Claude Aslangul

Mécanique quantique 2

Développements et applications à basse énergie

physique

COURS ET EXERCICES CORRIGÉS

LMD

Master de physique (M1)





Mécanique quantique 2

Licence Maîtrise Doctorat

Chimie

CACHAU-HERREILLAT D., *Des expériences de la famille Acide-Base.* 2e éd.
CACHAU-HERREILLAT D., *Des expériences de la famille Red-Ox*
DEPOVERE P., *Chimie générale.* 3e éd.
DEPOVERE P., *Chimie organique.* 2e éd.
MCMURRY J., BEGLEY T., *Chimie organique des processus biologiques*
MOUSSARD C., *Biochimie structurale et métabolique.* 3e éd.
MOUSSARD C., *Biologie moléculaire et Biochimie des communications cellulaires*
MOUSSARD C., GIBEY R., BÉNÉDINI M., *QCM de biochimie et de biologie moléculaire*
RABASSO N., *Chimie organique. Généralités, études des grandes fonctions et méthodes spectroscopiques*
RABASSO N., *Chimie organique. Hétéroéléments, stratégies de synthèse et chimie organométallique*

Biologie

FORÊT R., *Réussir le CAPES externe SVT*
TANZARELLA S., *Perception et communication chez les animaux*
THOMAS F., GUÉGAN J.-FR., RENAUD F., *Écologie et évolution des systèmes parasités*
GUÉGAN J.-F. ET CHOISY M., *Introduction à l'épidémiologie intégrative des maladies infectieuses et parasitaires*

Mathématiques

BOGAERT P., *Probabilités pour scientifiques et ingénieurs*
COTTET-EMARD F., *Analyse*
COTTET-EMARD F., *Analyse 2*
COTTET-EMARD F., *Algèbre linéaire et bilinéaire*
COTTET-EMARD F., *Calcul différentiel et intégral*
ÉTIENNE D., *Exercices corrigés d'algèbre linéaire. Tome 1*
ÉTIENNE D., *Exercices corrigés d'algèbre linéaire. Tome 2*
MARCHAND M., *Outils mathématiques pour l'informaticien.* 2e éd.

Physique

ASLANGUL C., *Mécanique quantique. 1. Fondements et premières applications*
BECHERRAWY T., *Optique géométrique*
BIEMONT É., *Spectroscopie atomique. Instrumentation et structures atomiques*
BIEMONT É., *Spectroscopie moléculaire. Structures moléculaires et analyse spectrale*
TAILLET R., *Optique physique. Propagation des ondes lumineuses*
WATSKY A., *Thermodynamique macroscopique*

Méthodologie

AIMONETTI J.-M., *Comment ne pas endormir son auditoire en 30 secondes. La communication orale avec diaporama*
POCHET B., *Méthodologie documentaire.* 2e éd.

Claude **Aslangul**

Mécanique quantique 2

Développements et applications à basse énergie

physique

COURS ET EXERCICES CORRIGÉS





Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web : **www.deboeck.com**

Éditions De Boeck Université
Rue des Minimes 39, B-1000 Bruxelles

Imprimé en Belgique

Dépôt légal :
Bibliothèque nationale, Paris : octobre 2008
Bibliothèque royale de Belgique : 2008/0074/260

ISBN 978-2-8041-5881-1

À Claire et Marie

Ce ne sont pas les perles qui font le collier,
c'est le fil.
(Gustave FLAUBERT)

Préface au tome II

Ce tome II propose des développements de la Mécanique quantique abordant des aspects à la fois conceptuels (la notion de symétrie, par exemple) et techniques (systèmes à plusieurs dimensions d'espace, seconde quantification) et procède dans le même état d'esprit que le tome I : malgré l'inévitable complexification du formalisme, exposer les idées fondamentales en prenant appui sur des considérations physiques simples et incontournables, et les formaliser dans le langage mathématique dont il est impossible de se passer. Toutefois, un autre objectif est de rendre la théorie quantique aussi accessible qu'elle le permet, recourant dans la mesure du possible à l'explication par les mots du langage courant, venant préparer ou illustrer les inévitables complications techniques. Au-delà du souci de transmettre au plus grand nombre, l'explication verbale, si elle ouvre souvent la voie à la compréhension en profondeur en permettant de reléguer (provisoirement !) le formalisme, fait aussi courir le risque, assumé, de ce que certains lecteurs avertis jugeront comme des redondances. Une (longue) expérience d'enseignement permet d'être convaincu que si, dans le meilleur des cas, les étudiants appréhendent la signification des termes d'une équation, ils éprouvent trop souvent de la difficulté quand il s'agit de leur donner réellement un sens.

La première partie de ce tome (chapitres 17 à 26) vise à développer le formalisme dans un cadre élargi, et à introduire au fur et à mesure les ajouts rendus nécessaires par la considération de situations physiques plus complexes. Le premier chapitre est consacré à la notion de symétrie[1] *et montre dans des situations très simples l'importance et la puissance des arguments fondés sur la nécessaire invariance des lois physiques quand un principe de relativité est admis. Il aborde également des symétries moins intuitives (invariance de jauge pour le champ électromagnétique, renversement du temps) mais dont l'importance justifie qu'en soit donné un exposé élémentaire dès que possible. Le chapitre suivant expose la théorie du moment cinétique, l'une de ces grandeurs physiques dont la nécessaire quantification est au cœur de la théorie quantique ; la composition*

[1]Pierre Curie fut l'un des premiers à relever l'importance de la notion de symétrie en Physique. En 1894, il écrivait [1] :

> *Lorsque certaines causes produisent certains effets, les éléments de symétries des causes doivent se retrouver dans les effets produits. [...] Lorsque certains effets révèlent une certaine dissymétrie, cette dissymétrie doit se retrouver dans les causes qui lui ont donné naissance.*

des moments cinétiques est abordée en détail, avant d'établir l'important théorème de Wigner - Eckart et de montrer ses toutes premières applications (règles de sélection).

Le chapitre 19 s'inscrit dans la lignée des deux précédents en traitant le cas du champ central, pour lequel la symétrie sphérique se traduit par la conservation du moment cinétique. Différents systèmes sont traités : particule libre, puits carré sphérique, atome d'hydrogène. Chacun d'entre eux est d'une part l'occasion de voir à l'œuvre les arguments de symétrie, d'autre part fournit un exemple de passage d'une représentation à l'autre (des ondes planes aux fonctions de Bessel sphériques, par exemple). En outre, c'est une belle occasion de rencontrer quelques fonctions spéciales apparaissant si souvent en Physique qu'une familiarité avec elles est nécessaire, même si elle doit rester à un niveau superficiel. On s'est efforcé de suivre la voie la plus directe pour rendre ces aspects techniques aussi élémentaires que possible, espérant d'ailleurs avoir contribué à démystifier des objets ou des notions dont le simple énoncé de leur appellation provoque parfois quelques raideurs chez l'auditeur non encore averti. Après tout, la fonction hypergéométrique – surtout quand elle est dégénérée – n'est pas la fonction de Cantor et n'a vraiment rien de redoutable ; quant à la fonction Gamma ou les fonctions de Bessel, elles sont si universelles en Physique qu'elles appartiennent sans conteste à la boîte à outils de base : sauf à faire le choix de l'amateurisme, on ne peut prétendre faire sérieusement de la Physique sans un bagage minimum de mathématiques appliquées – ce qui était vrai dans la bouche de Galilée l'est encore plus quand il s'agit de manipuler avec aisance la théorie quantique. N'est pas Feynman qui veut, lequel n'ignorait rien des mathématiques les plus avancées – n'invente pas l'intégrale de Feynman qui veut !

Le chapitre 20 est une introduction physique au spin, que l'on peut considérer comme un degré de liberté spécifiquement quantique. Il y est montré comment l'existence avérée du moment magnétique propre de l'électron oblige à attribuer à l'électron un moment cinétique propre (spin) afin de résoudre une difficulté physique majeure. Ce chapitre se termine par une présentation de l'équation de Dirac procédant par essai et par erreur, afin de montrer comment une équation historiquement fondamentale peut graduellement s'imposer sur la base de considérations à la fois techniques et heuristiques. Les difficultés de fond de cette théorie sont discutées, et fournissent un nouvel exemple des progrès rendus possibles quand, parvenue dans une impasse, la théorie doit être reprise sur de nouvelles bases ; si la théorie de Dirac est parfois considérée comme obsolète, il convient de retenir que c'est son échec, relatif mais indiscutable, qui a permis l'essor de la Théorie quantique des champs.

Une fois exposés ces éléments conceptuels et techniques, il est possible de revenir sur les fondements de la Mécanique quantique et de décrire des expériences assez récentes faisant passer des expériences de pensée imaginées par les Pères-fondateurs aux réalités de laboratoire ; c'est l'objet du chapitre 21. L'effet Zénon y est discuté formellement, avant la description et l'interprétation de l'expérience l'ayant mis pour la première fois en évidence. Il en va de même pour les sauts quantiques, dont Schrödinger, le premier, doutait fort qu'ils puissent exister. Les bases de la cryptographie quantique sont données, l'une des applications sans doute les plus inattendues de la Théorie quantique. Finalement, des notions élémentaires sur la théorie de la décohérence sont proposées, théorie visant à une meilleure compréhension du pourquoi de l'apparence classique du monde

perceptible avec le sens commun.

On sait que l'une des renonciations obligées dans le cadre quantique est celle du concept de trajectoire. Dans le chapitre 22, on montre comment cette nécessité conduit à de nouvelles difficultés lorsque le système est composé de particules identiques, ayant par définition les mêmes attributs intrinsèques (masse, charge, spin,...). L'impossibilité de suivre chacune sur une trajectoire qui n'existe pas produit une indiscernabilité *qui, sans l'énoncé d'un nouveau postulat, rendrait impossible la prévision de certains résultats d'expérience. Le postulat supplémentaire donne lieu à la définition de deux grandes classes de particules (fermions et bosons), aux propriétés pour le moins contrastées. Il introduit aussi une forme de corrélation très subtile entre des particules indépendantes au sens classique et est le fondement de l'explication ultime, tant sur le plan qualitatif que quantitatif, de phénomènes majeurs autrement incompréhensibles comme le magnétisme atomique, la distinction isolant/conducteur ou la supraconductivité. La complexité du formalisme primitif de la théorie quantique a imposé une autre écriture plus maniable, dite Seconde quantification, dont les bases sont expliquées sommairement.*

L'un des buts du physicien est la construction de modèles *visant à saisir l'essence des phénomènes ; même ultra-simplifiés, les modèles ainsi définis sont rarement solubles exactement. La Mécanique quantique n'échappe donc pas à la nécessité de pourvoir des méthodes d'approximation destinées à rendre possible leur exploitation. Le chapitre 23 est consacré à deux grands schémas : la méthode dite variationnelle, et la théorie des perturbations stationnaires, permettant d'obtenir des approximations pour les états propres. Il ne s'agit pas de dresser le catalogue des nombreuses variantes disponibles dans la littérature mais de donner les principes d'approximations autorisant à construire un schéma plus spécifique et mieux adapté à tout problème concret examiné.*

Le chapitre suivant expose la théorie des perturbations dépendant du temps, à la base de la compréhension de l'extrême variabilité de l'intensité des raies atomiques par le jeu des règles de sélection. La Règle d'or de Fermi est démontrée et permet la description quantique de l'un des phénomènes majeurs rencontrés dans le tome I : l'effet photoélectrique. Initialement édifiée pour décrire l'évolution d'un système couplé à un champ classique variable dans le temps (un atome soumis au rayonnement électromagnétique, par exemple), la théorie des perturbations dépendant du temps s'est révélée porteuse d'une grande fécondité pour traiter des problèmes statiques, notamment dans le problème à N-corps, dont un exemple est l'important théorème de Gell-Man et Low. Il est aussi montré comment la théorie perturbative peut être très largement dépassée en effectuant des resommations partielles conduisant notamment au calcul simplifié de la durée de vie et du déplacement de Lamb d'un atome.

Le chapitre 25 est une introduction à la quantification du champ électromagnétique, permettant non seulement de formaliser un phénomène aussi important que l'émission spontanée, mais aussi de revisiter des notions rencontrées à un niveau plus élémentaire, comme les forces de van der Waals entre deux objets neutres électriquement. On montre notamment comment cette situation fournit un bel exemple d'interaction par échange de particules virtuelles (ici des photons).

Le chapitre 26 est une ouverture sur la théorie de la diffusion, dont l'importance pratique peut se mesurer au fait qu'un très grand nombre d'expériences consiste à "éclairer" une cible, l'analyse des caractéristiques des produits finals permettant de remonter aux interactions élémentaires. Techniquement subtiles, ne serait-ce que par la nécessité d'impliquer des états de continuum, *les deux approches conventionnelles (équation intégrale et méthode des déphasages) sont présentées de façon aussi simple que possible, tout en dégageant des notions universelles, comme celle de fonction de Green.*

La deuxième partie de l'ouvrage (chapitres 27 à 32) vise à susciter la curiosité du lecteur en proposant un florilège de quelques applications de la Mécanique quantique en Physique de basse énergie. Le chapitre 27 donne les idées de base pour la description des atomes, systèmes devant déjà faire l'objet de traitements approchés puisqu'ils constituent l'exemple le plus simple de problème à N-corps. Les notions de structures fine et hyperfine, si importantes pour les tests de l'Électrodynamique quantique ou en Astrophysique, sont présentées. Le chapitre suivant propose un tour d'horizon de la description des molécules ; après une présentation très intuitive de l'approximation de Born et Oppenheimer, les fondements physiques de la liaison chimique sont présentés, puis on jette les bases de la spectroscopie moléculaire, tant dans le domaine UV que dans les domaines de l'infra-rouge et des microondes.

Les trois derniers chapitres abordent la matière condensée, un monde en soi, délibérément limité aux solides cristallins. On présente d'abord de façon semi-quantitative les clés permettant de comprendre l'existence des différents types de solides, avant de montrer comment l'ordre spatial à longue portée donne lieu à la diffusion sélective de Bragg, fondement de méthodes expérimentales incontournables pour la connaissance précise de la structure des solides.

Le chapitre 30 commence par montrer l'importance de la symétrie discrète de translation, et sa conséquence immédiate (le théorème de Bloch), qui donne au mouvement d'un électron dans un cristal des propriétés très particulières ; alliées au Principe de Pauli, elles débouchent sur l'un des succès spectaculaires de la Mécanique quantique : la compréhension du pourquoi de la conductivité électrique et donc de la distinction isolant/conducteur. Sont ensuite présentées quelques propriétés magnétiques, avec notamment l'introduction des magnons, *superbe exemple de quasi-particules pouvant former des paires liées.*

Les vibrations de réseau sont abordées dans le chapitre suivant. On montre comment leur quantification sous forme de phonons permet de rendre compte du comportement basse température de la chaleur spécifique des solides, autre conséquence de l'hypothèse révolutionnaire de Planck, comme l'a de suite montré Einstein. En outre, le retour sur la diffraction en présence de phonons donne l'occasion de rencontrer un exemple où la dimensionnalité des systèmes joue un rôle important.

Le dernier chapitre se veut d'abord une initiation au problème du transport dans les solides, et montre notamment comment il est possible de calculer la conductivité d'un métal, faisant naturellement émerger le concept de masse effective, *à la base de patientes recherches en physique des matériaux. L'ouvrage se termine par une initiation à la*

supraconductivité, et présente les fondements de l'une des plus belles théories construites dans un cadre quantique : la théorie BCS.

L'esprit et la présentation de ce second tome sont identiques à ceux du premier. On n'hésite pas à faire mention ici et là de résultats importants très au-delà du niveau M1, dont le seul but est une ouverture sur une question assez pointue visant à piquer la curiosité du lecteur, et l'invitant implicitement à consulter des ouvrages très spécialisés, ou à dénicher un (bon) site web permettant d'en apprendre plus. Il est souvent fait appel, par des analogies et des rapprochements, aux connaissances acquises au même niveau dans d'autres secteurs de la Physique : si la démarche se veut essentiellement déductive, elle n'est pas forcément toujours linéaire et s'autorise des diversions quand le point en discussion est l'occasion de montrer l'universalité d'une méthode ou d'un raisonnement utile dans des spécialités trop souvent disjointes : le polyacétylène est un bel exemple prouvant, s'il le fallait, que les préoccupations du physicien et du chimiste se rejoignent et s'enrichissent mutuellement.

Chaque chapitre se termine par une section d'exercices et de problèmes s'efforçant de réaliser un équilibre entre la familiarisation avec des idées physiques et le développement de l'habileté technique indispensable. La difficulté des problèmes est très variable, allant de simples applications de cours à des questions ayant fait l'objet d'articles de recherche récents, en passant par un prétexte à la familiarisation de notions importantes écartées faute de place (la théorie des groupes, les méthodes de fonction de Green, par exemple). Des références, éventuellement à des articles originaux, permettent au lecteur intéressé d'aller plus loin dans l'analyse d'un point lui paraissant digne d'intérêt. Enfin, des corrigés sont disponibles à l'adresse `http://universite.deboeck.com/`.

Le choix des références est toujours difficile. On trouvera celles des articles pionniers, contrastant avec des articles de synthèse récents pour les sujets ayant donné lieu à une abondante littérature. Par ailleurs, il sera souvent fait référence aux magnifiques cours de Claude Cohen-Tannoudji donnés au Collège de France, dont les notes sont disponibles sur la Toile [2] et qui, pour certains, ont été édités [3]. Au-delà de la citation obligatoire, l'auteur souhaite qu'y soit vue également l'expression de la reconnaissance d'avoir tant pu apprendre durant des décennies.

Remerciements

Inscrit dans la continuité du premier, ce deuxième tome est grandement redevable à celles et ceux dont l'attention a forgé l'heureux devoir pour moi de leur exprimer ma gratitude. La proximité quotidienne de Dominique Mouhanna a été particulièrement salutaire, tout comme les discussions avec Bertrand Delamotte et Jean-Marie Maillard. Confronté à devoir présenter des questions au départ peu familières pour moi, leur patience à m'expliquer des sujets qu'ils maîtrisent pour y avoir si profondément réfléchi a été d'un très grand secours.

Issus d'un enseignement à l'E.N.S. et à l'Université Pierre et Marie Curie, ces ouvrages ont grandement bénéficié de discussions critiques et éclairantes avec les étudiants les plus enthousiastes et les plus soucieux de comprendre en profondeur. Je remercie à nouveau Fabien Beckers, Eli Ben-Haïm, Gaëtan Borot, Alexandre Céalis, Olivier Deloubrière, Alexandre Flavier, Céline Laroche, Juliette Réallomble et Julien Vidal pour leurs questions pertinentes et leurs remarques incisives, et pour leur enthousiasme vivifiant.

En me donnant la possibilité d'enseigner en première année à l'E.N.S., Jean-Michel Raimond m'a conduit à devoir approfondir des questions le plus souvent absentes des ouvrages universitaires, malgré leur importance et la démonstration qu'elles apportent de la profondeur, de la subtilité et de la robustesse de la théorie quantique. Je lui sais gré de la confiance ainsi accordée.

Tout au long de leur élaboration, Monsieur Francis Germain aura veillé sur ces ouvrages avec un sens aigu de l'analyse. Son insatiable curiosité intellectuelle aura été un aiguillon permettant d'améliorer grandement des points dont la subtilité exige une progression par étapes et une judicieuse articulation des idées. Qu'il soit remercié pour sa persévérance, qui aura aussi permis de corriger des erreurs dans le texte préliminaire.

Je suis profondément redevable à Mr. Fabrice Chrétien, des Éditions De Boeck, d'avoir pris l'initiative de ce projet éditorial, donnant l'occasion du larron de découvrir une activité passionnante et enrichissante. Mme Florence Lemoine a relu avec un soin admirable le document final : qu'elle en soit vivement et chaleureusement remerciée.

Parce qu'elles m'ont permis de jouer pendant des années le rôle d'un précepteur quelque peu archaïque, Claire et Marie ont développé la volonté de transmettre ce que l'on sait, animées par le désir d'apprendre et de comprendre ce qui, en apparence, ne servira jamais à rien : qu'elles soient remerciées pour avoir si grandement contribué au bonheur de partager un peu de vrai savoir. Enfin, mes derniers mots seront pour Celle qui aura accompagné des mois durant ces centaines de pages au point de pouvoir dire, en métaphore lumineuse, que ces livres furent écrits à deux cœurs et à quatre mains.

Table des Matières

IV Applications à basse énergie 1221

Partie III

Développements

Chapitre 17

Symétrie et lois de conservation

Le but de ce chapitre est de montrer explicitement la relation entre la symétrie, le plus souvent géométrique, et les lois de conservation qui en découlent. Le lien est établi d'abord en Mécanique Classique, puis en Mécanique quantique, mettant ainsi en évidence les parallèles entre ces deux théories.

17.1 Le principe euclidien de relativité et le rôle de la symétrie en Physique

L'espace où se situent et se déplacent les systèmes physiques est supposé doué de propriétés géométriques permettant de donner un sens à la comparaison d'expériences accomplies avant ou après avoir effectué une certaine opération (translation, rotation) modifiant la position d'un point dans l'espace, telle qu'elle est définie par trois cordonnées relativement à un repère donné. En Physique non-relativiste au sens d'Einstein, où la vitesse de la lumière c est formellement infinie, le paramètre temps est découplé des coordonnées spatiales : dans l'espace-temps à quatre dimensions $\mathbb{R}^4$, tous les points physiques se trouvent à tout instant t_0 dans le *même* (hyper-)plan $t = t_0$. Il en résulte qu'un point matériel classique évolue le long d'une trajectoire de $\mathbb{R}^3$, paramétrée par le temps t, qui est le même pour tous les observateurs. La distance de deux points de l'espace-temps à un instant donné se réduit alors à la distance euclidienne ; relativement à un repère cartésien Oxyz, elle vaut, par définition :

$$\boxed{d^2 \stackrel{\text{déf}}{=} x^2 + y^2 + z^2} \qquad (17.1)$$

En outre, l'espace est supposé posséder les mêmes propriétés en deux points se déduisant l'un de l'autre par une rotation quelconque (*isotropie* de l'espace), et en deux points se déduisant l'un de l'autre par une translation également quelconque (*uniformité*, ou *homogénéité*, de l'espace). L'ensemble de ces opérations constitue visiblement un groupe (groupe d'Euclide), à 6 paramètres : trois permettent de définir toute translation, et il en faut autant pour définir toute rotation[1] (deux angles pour fixer l'axe de la rotation, un troisième pour l'angle). Le groupe d'Euclide ne contient pas le paramètre *temps* et n'inclut de ce fait aucun élément de nature *cinématique*.

Qu'elle soit de nature théorique ou expérimentale, l'étude des phénomènes physiques requiert l'introduction du temps, et l'usage d'un système de référence (repère) par rapport auquel on définit et/ou mesure les attributs d'un système (sa position, sa vitesse...). D'un repère à l'autre, les lois déterminant le mouvement n'ont pas en général la même forme (penser à un repère accéléré, rectilignement ou en rotation – la position d'équilibre d'un pendule simple dans un train qui freine n'est pas la verticale). S'en tenant à la Dynamique, c'est-à-dire à la description du mouvement des corps sous l'effet des forces qui leur sont appliquées, l'équation fondamentale est celle dite de Newton :

$$\boxed{m\frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = \vec{F}} \tag{17.2}$$

L'hypothèse fondamentale de la Physique selon Galilée[2] est un principe suivant lequel l'espace ordinaire (physique) est isotrope et homogène et où le temps est le même partout (*uniformité* du temps) : un système isolé et fermé étant défini, ses propriétés ne changent pas si on le déplace d'un point à l'autre[3] ou si on le fait tourner sur lui-même ; une même expérience doit donner le même résultat, qu'elle soit démarrée à midi ou à l'heure du goûter. Si les propriétés sont inchangées, aucun résultat d'expérience ne doit être affecté par de telles opérations, et c'est pourquoi toute théorie physique visant à décrire les propriétés des systèmes doit posséder des propriétés d'invariance : *les lois qu'elle énonce doivent être les mêmes dans une classe d'équivalence de repères remarquables* ; cette affirmation est à la fois une déclaration de principe de l'existence de ces repères, appelés *galiléens* (ou encore *référentiels d'inertie*), et leur définition. Dans un tel repère, par définition même, l'équation de Newton (17.2) est vraie.

De façon remarquable, l'équation de Newton (17.2) n'est pas seulement invariante dans les opérations du groupe d'Euclide[4], elle l'est aussi dans toute transformation

[1]Dans cette discussion, on ne considère que les groupes continus (dits parfois *connexes*, par opposition aux groupes ne contenant que des opérations discrètes (voir section 17.5).

[2]Voir [4], p. 371. Merzbacher emploie l'expression *Euclidean Principle of Relativity*, usage qui se justifie tant les concepts euclidiens sont finalement très voisins de ceux de Galilée. Pour cette raison, on s'autorisera dans la suite à utiliser les adjectifs *euclidien* et *galiléen* en tant que synonymes.

Le mot *relativité* renvoie le plus souvent à la relativité au sens d'Einstein, et à sa formulation première en 1905. En réalité, cette notion est présente en Physique depuis Galilée. La relativité galiléenne s'obtient à partir de celle d'Einstein en faisant formellement $c = +\infty$, mais le concept lui-même demeure intact et ne disparaît pas de la Physique "non-relativiste". Sur les développements historiques de cette notion, voir l'ouvrage fondamental de M.A. Tonnelat [5].

[3]Cette affirmation contient implicitement l'hypothèse selon laquelle l'espace est infini dans toutes les directions.

[4]Comme dit plus haut, on ne considère ici que les symétries continues. L'équation (17.2) a également

échangeant deux repères en translation uniforme (à vitesse constante) l'un par rapport à l'autre. C'est cette invariance qui exprime très précisément la relativité au sens de Galilée[5] (voir l'encadré p. 659), et conduit à l'introduction du groupe de Galilée, ensemble des opérations reliant entre eux tous les référentiels d'inertie[6].

Dans l'ouvrage *Dialogue sur les deux plus grands systèmes du monde*[a] (1632), Galilée écrit :

> *Enfermez-vous avec un ami dans la cabine principale à l'intérieur d'un grand bateau et prenez avec vous des mouches, des papillons, et d'autres petits animaux volants. Prenez une grande cuve d'eau avec un poisson dedans, suspendez une bouteille qui se vide goutte à goutte dans un grand récipient au-dessous d'elle. Avec le bateau à l'arrêt, observez soigneusement comment les petits animaux volent à des vitesses égales vers tous les côtés de la cabine. Le poisson nage indifféremment dans toutes les directions, les gouttes tombent dans le récipient au-dessous, et si vous lancez quelque chose à votre ami, vous n'avez pas besoin de le lancer plus fort dans une direction que dans une autre, les distances étant égales, et si vous sautez à pieds joints, vous franchissez des distances égales dans toutes les directions. Lorsque vous aurez observé toutes ces choses soigneusement (bien qu'il n'y ait aucun doute que lorsque le bateau est à l'arrêt, les choses doivent se passer ainsi), faites avancer le bateau à l'allure qui vous plaira, pour autant que la vitesse soit uniforme et ne fluctue pas de part et d'autre. Vous ne verrez pas le moindre changement dans aucun des effets mentionnés et même aucun d'eux ne vous permettra de dire si le bateau est en mouvement ou à l'arrêt ...*

[a] *Dialogo sopra i due massimi sistemi del mondo.* Ce dialogue se tient entre trois interlocuteurs : Salviati, partisan de Copernic, Sagredo, un vénitien éclairé et ouvert, et Simplicio, défenseur simpliste (!) de la pensée d'Aristote et du géocentrisme.

Sur un plan technique, la relativité galiléenne s'exprime comme suit. Soit un point matériel se déplaçant dans $\mathbb{R}^3$, dont les coordonnées à l'instant t sont $x(t)$, $y(t)$, $z(t)$ par rapport à un certain repère R trirectangle $\mathrm{O}xyz$. Soit un autre repère R', $\mathrm{O}'x'y'z'$, se déplaçant à la vitesse constante $\vec{V}$ par rapport à R, le vecteur $\vec{V}$ étant parallèle à $\mathrm{O}x$, et les trois axes $\mathrm{O}'x'$, $\mathrm{O}'y'$ et $\mathrm{O}'z'$ étant respectivement parallèles à $\mathrm{O}x$, $\mathrm{O}y$ et $\mathrm{O}z$. Par rapport à R', les coordonnées du même point matériel à l'instant t' sont $x'(t')$, $y'(t')$, $z'(t')$, et sont

des symétries discrètes (par exemple, invariance par renversement du temps), voir section 17.5.

En outre, cette équation suppose implicitement l'existence d'une *trajectoire*, au sens d'une ligne continue possédant une tangente, au moins par morceaux. Les courbes rugueuses à toute échelle (lignes continues sans tangente) sont donc d'emblée exclues dès que l'équation de Newton est posée. On sait que la notion de trajectoire n'est pas compatible avec les concepts quantiques, comme en témoigne de façon très condensée la relation $\Delta x \Delta p_x \geq \frac{\hbar}{2}$.

[5] Galileo Galilei (1564-1642) commença des études de médecine, sans y porter beaucoup d'intérêt et sans obtenir de diplôme. Son œuvre de physicien est considérable, tant par les expériences que son habileté manuelle lui permettait d'accomplir que par l'affirmation définitive de concepts qui avaient valu à Giordano Bruno d'être brûlé vif quelques décennies auparavant. Si le sort de Galilée fut moins tragique, son *Dialogue* servit de chef d'accusation au procès qui le condamna à la prison à vie, à l'issue duquel il aurait, selon la légende, prononcé le fameux *E pur si muove!*.

Son père, Vincenzo, était luthier et musicien, et a laissé quelques œuvres charmantes encore inscrites au répertoire des instrumentistes d'aujourd'hui.

[6] Ce groupe est justement la limite $c = +\infty$ du groupe de Lorentz sous-jacent à la Relativité au sens d'Einstein.

reliées aux coordonnées dans R par les formules suivantes (transformation de Galilée) :

$$\boxed{t' = t \ , \quad x' = x - Vt \ , \quad y' = y \ , \quad z' = z} \tag{17.3}$$

Le groupe de Galilée pour trois repères R, R' et R" est *ipso facto* défini par la règle de composition des vitesses $\vec{V}'' = \vec{V} + \vec{V}'$, qui résulte directement de (17.3). Le produit de deux transformations de Galilée est visiblement une transformation de Galilée.

Bien évidemment, ce principe s'étend au cas d'un système en interaction avec un autre, ou avec son milieu : pour un système soumis à un champ de forces variable ou non dans le temps, les propriétés ne changent pas si le champ de forces est déplacé ou tourné en même temps que le système pré-défini. En réalité, l'ensemble de ce dernier et le seul champ de forces supposé exister constitue un système isolé et fermé auquel le principe de relativité s'applique.

Ainsi, le principe de relativité de Galilée nie l'existence d'une position ou d'une orientation privilégiée dans l'espace où sont plongés les objets que la Physique prétend étudier et décrire : l'espace est déclaré uniforme et isotrope, et le paramètre temps est universel[7]. Un corollaire immédiat de ce principe[8] est qu'il n'existe pas de repère "absolu" à privilégier, toutes les lois de la Nature devant être les mêmes dans tous les référentiels galiléens concevables, où qu'ils se trouvent dans l'univers. C'est la violation avérée de cette affirmation pour l'Électromagnétisme[9] qui a conduit Lorentz[10] à redéfinir la transformation devant relier deux repères universels, avant l'écriture par Einstein (1905) de la Relativité restreinte.

[7]Toutes ces affirmations reposent également sur un non-dit : un point de l'espace-temps est caractérisable par des grandeurs *continues*. À la fin de sa vie, Schrödinger [6] a attiré l'attention sur le fait qu'une telle *évidence* mériterait d'être reconsidérée.

[8]L'affirmation contenue dans le principe de relativité de Galilée ne requiert par nature aucune démonstration : ériger en principe une affirmation, c'est s'en remettre pour la preuve de sa véracité à la confrontation entre les résultats théoriques qui en découlent et les observations expérimentales.

[9]L'insuffisance de la transformation de Galilée devint patente plusieurs décennies après l'expérience de Michelson et Morley (1887) destinée à mettre en évidence le mouvement de la Terre par rapport à l'Éther.

Albert Abraham Michelson (1852-1931) reçut le Nobel en 1907 pour *"for his optical precision instruments and the spectroscopic and metrological investigations carried out with their aid"*. Edward Morley (1838-1923) fut professeur de Chimie à Case Western Reserve University.

La mention du Comité Nobel et l'"oubli" de Morley montrent combien le résultat négatif de leur expérience ne fut pas de suite perçu comme le point de départ d'une révolution conceptuelle. Certains, d'ailleurs, affirment que ni Michelson, ni Morley ne considéraient l'issue de leur expérience comme apportant la preuve de l'inexistence de l'Éther. Pour un exposé détaillé, voir l'article de Fowler [7] ; les différentes interprétations sont données dans le livre de Simon [8].

[10]Hendrik Antoon Lorentz (1853-1928) présenta une explication de l'expérience de Michelson et Morley sur la base d'une réelle contraction des longueurs (1904). Il avait partagé le Nobel en 1902 avec Pieter Zeeman *"in recognition of the extraordinary service they rendered by their researches into the influence of magnetism upon radiation phenomena"*, pour leur découverte de l'effet portant le nom de ce dernier.

Les travaux de Lorentz furent en un sens confortés par ceux de Henri Poincaré (1854-1912), mais ni l'un ni l'autre ne remirent en cause l'existence d'un temps et d'un Éther absolus. Outre ses travaux qui ont indiscutablement ouvert la voie à Eisntein (comme en témoigne leur correspondance), Poincaré est universellement connu pour ses contributions décisives à des problèmes très divers (optique, problème des trois corps, équations différentielles et systèmes dynamiques – ce qui ne s'appelait pas encore la théorie du chaos). Son ouvrage *La Science et l'hypothèse* [9] reste une référence en épistémologie, dont la lecture est vivement recommandée.

En Physique, on nomme habituellement "le système" un ensemble d'objets sur lesquels le "reste du monde" (environnement) peut agir ; à l'inverse, cet ensemble n'a pas de rétroaction sur l'environnement. Dire qu'une particule est soumise à un champ (électrique, magnétique, électromagnétique) signifie que les sources de ce champ (condensateur, aimant, courants variables dans le temps) sont supposées avoir leur dynamique propre définie en soi, quoi que fasse la particule – qui constitue alors le système physique d'étude. Il est clair que ceci, dans l'absolu, est toujours une approximation à justifier, tout comme le choix d'un repère déclaré galiléen : construire la Mécanique quantique dans le cadre du principe de Relativité a un sens, même vis-à-vis des expériences faites sur Terre ; s'il y existe évidemment une direction privilégiée (le sens du champ de force de gravitation), ses effets, comparés à ceux qui sont pertinents pour la dynamique des petits objets (particules, atomes, molécules, etc.), sont totalement négligeables, sauf cas très exceptionnel[11].

À titre d'exemple, montrons comment le principe de relativité permet[12] de construire le Lagrangien L d'une particule libre, c'est-à-dire soumise à aucune force. L'uniformité (homogénéité) de l'espace autorise à affirmer que L ne dépend d'aucune coordonnée spatiale, puisque toutes les positions sont équivalentes (la particule "ne sait pas où elle est"). De la même façon, L ne peut dépendre du temps puisque tous les instants se valent. Il en résulte que L est au plus une fonction de la vitesse. Utilisons maintenant l'isotropie de l'espace ; comme toutes les directions sont équivalentes, L ne peut dépendre de l'orientation du vecteur vitesse : seul importe le *module* de ce vecteur. Celui-ci pouvant toujours s'exprimer en fonction du carré scalaire $\vec{v}^2 \equiv v^2$, on en tire la conclusion que le Lagrangien d'une particule libre, tel qu'il résulte du principe de Galilée, est nécessairement une certaine fonction du seul carré de la vitesse :

$$\boxed{L = f(v^2)} \tag{17.4}$$

la fonction f restant à déterminer.

Le Lagrangien est au cœur de l'intégrale donnant l'action, S, qui est extrémale (minimale, c'est le Principe de Moindre Action, voir Tome I, chapitre 7) pour la trajectoire réellement suivie :

$$S = \int_{t_1}^{t_2} L\,\mathrm{d}t \ , \qquad \delta S = 0 \text{ pour la trajectoire réellement suivie} \ . \tag{17.5}$$

Soit un point matériel (particule) libre, dont les coordonnées dans deux repères galiléens R et R' sont $\vec{r}$ et $\vec{r}\,'$, R' étant animé d'une vitesse $\vec{V}$ par rapport à R ; selon (17.3), on a :

$$\vec{r}\,' = \vec{r} - \vec{V}t \ , \qquad \vec{v}\,' = \vec{v} - \vec{V} \ . \tag{17.6}$$

Dans chacun de ces repères, le mouvement est décrit par les équations de Lagrange déduites du principe de moindre action. Soit L et L' les deux Lagrangiens manipulés

[11] La description très précise de la dynamique de particules ultra-froides (neutrons, atomes, etc.) peut exiger l'incorporation du champ de pesanteur terrestre.

[12] voir [10], p. 12.

par deux observateurs appartenant l'un à R, l'autre à R'. Comme pour ces deux observateurs les lois du mouvement doivent être les mêmes, les deux actions S et S' ne doivent différer que d'une constante, que la procédure de minimisation fera disparaître ; il en ira bien ainsi à condition que la différence $L - L'$ soit une dérivée totale en temps, du genre $\frac{\mathrm{d}}{\mathrm{d}t}\phi(t)$: alors, l'intégration en temps donnant l'action produira $\phi(t_2) - \phi(t_1)$, une constante s'éliminant d'elle-même lors de la variation ultérieure δS conduisant aux équations de Lagrange.

Plus précisément, les deux observateurs vont respectivement écrire :

$$\delta \int_{t_1}^{t_2} L\,\mathrm{d}t = 0 \ , \qquad \delta \int_{t_1}^{t_2} L'\,\mathrm{d}t = 0 \ . \tag{17.7}$$

Le premier observateur en déduira une loi exprimée par :

$$\boxed{\vec{r} = g(t,\, t_0\,;\, \vec{r}_0, \vec{v}_0)} \tag{17.8}$$

où g est une fonction du temps paramétrée par les conditions initiales $\vec{r}_0, \vec{v}_0$. Le deuxième observateur trouvera la *même* loi, exprimée par la *même* fonction g mais paramétrée par les conditions initiales mesurées dans R' :

$$\boxed{\vec{r}\,' = g(t,\, t_0\,;\, \vec{r}\,'_0, \vec{v}\,'_0)} \tag{17.9}$$

le temps étant le même dans les deux référentiels. Écrivons maintenant que l'identité des lois exige que L et L' diffèrent seulement par une dérivée totale en temps. Partant de :

$$L = f(v^2) \ , \qquad L' = f({v'}^2) \tag{17.10}$$

et prenant en compte (17.6), L' s'écrit :

$$L' = f(v^2 - 2\vec{v}.\vec{V} + V^2) \ . \tag{17.11}$$

Supposons que la vitesse $\vec{V}$ est un infiniment petit ; en effectuant le développement de Taylor du second membre, on obtient L' à des infiniment petits d'ordre 2 près :

$$L' = f(v^2) \ - 2\vec{v}.\vec{V}\,\frac{\partial f}{\partial v^2} + \mathcal{O}(v^2) \equiv L \ - 2\vec{v}.\vec{V}\,\frac{\partial f}{\partial v^2} + \mathcal{O}(v^2) \ . \tag{17.12}$$

Le terme additionnel doit être de la forme $(\mathrm{d}/\mathrm{d}t)\phi(t)$ pour que les lois trouvées dans R et R', déduites de L et de L', coïncident. Comme $\vec{v} \equiv (\mathrm{d}/\mathrm{d}t)\,\vec{r}$ est déjà une dérivée totale, le terme complémentaire s'écrit $-2\vec{V}.\frac{\mathrm{d}\vec{r}}{\mathrm{d}t}\,\frac{\partial f}{\partial v^2}$; il ne peut être de la forme souhaitée que si $\frac{\partial f}{\partial v^2}$ est une constante[13] : f est donc un monôme du genre av^2 où a désigne une constante, $f(v^2) = av^2$, donnant finalement :

$$\boxed{L = av^2} \tag{17.13}$$

La constante a n'est autre que la moitié de la masse m de la particule – ceci afin de retrouver l'équation de Newton. Une fois obtenue cette dépendance pour une transformation

[13] Penser à $(fg)' = f'g + fg'$.

infinitésimale, il est facile de voir que le Lagrangien $L = av^2$ est également invariant[14] pour une transformation finie :

$$L' = av'^2 = a(\vec{v} - \vec{V})^2 = a(v^2 - 2\vec{v}.\vec{V} + V^2) \equiv L + a\,\frac{\mathrm{d}}{\mathrm{d}t}\,(-2\vec{r}.\vec{V} + V^2 t) \ , \qquad (17.14)$$

qui est bien de la forme :

$$\boxed{L' = L + \frac{\mathrm{d}}{\mathrm{d}t}\,\phi} \qquad (17.15)$$

Comme les deux lagrangiens L et L' ne diffèrent que par une dérivée totale en temps, ils donnent les mêmes équations du mouvement[15] dans les deux repères, que la transformation de R à R' soit infinitésimale ou non.

Pour une particule soumise à un champ (de gravitation, électrostatique, etc.) dérivant d'un potentiel U, le Lagrangien s'écrit[16] :

$$\boxed{L = T - V} \qquad (17.16)$$

où V est l'énergie potentielle, différant simplement du potentiel U par un scalaire multiplicatif[17] (masse pour le champ de gravitation, charge pour le champ électrostatique, etc.)[18]. Le choix exprimé par (17.16) assure que les équations de Lagrange déduites du principe variationnel $\delta S = 0$ reproduisent les équations de la Mécanique dans sa formulation élémentaire.

L'énergie potentielle V est *a priori* une fonction des trois coordonnées d'espace x, y et z, ou de tout autre jeu de coordonnées fixant sans ambiguïté la position de la particule dans l'espace. V est la représentation désincarnée physiquement de l'action de l'extérieur sur la particule (celle-ci étant le système d'étude) : on a remplacé tous les autres systèmes en interaction avec la particule par un champ de forces "éthérées" ; ces forces agissent sur la particule, qui n'a en retour aucune influence sur elles. C'est ainsi que s'introduit, à un niveau élémentaire, la notion de *champ* en Physique.

La symétrie proclamée par le principe de Galilée pour l'espace où plongent les systèmes physiques est une symétrie *universelle*. De surcroît, un système donné, dans un environnement donné, peut posséder une symétrie qui lui est propre, que l'on peut

[14]Au sens : produit des équations du mouvement invariantes.

[15]On retrouvera plus loin cet argument à propos de l'invariance de jauge (voir Remarque p. 689). Ceci permet de considérer la transformation de Galilée comme un cas particulier de transformation de jauge au sens le plus large (pas seulement définie à propos du champ électromagnétique).

[16]Voir Tome I, chapitre 7.

[17]Voir la note 19 p. 664 pour le champ de gravitation terrestre.

[18]La formulation de la Mécanique quantique est fondée sur la description hamiltonienne habituelle, où il est possible de définir une énergie potentielle V. On peut toujours se trouver dans cette situation en élargissant suffisamment le système à quantifier – et en ne considérant d'emblée que des interactions à deux corps –, quitte à prendre des traces partielles par la suite si seule une partie des degrés de liberté est digne d'intérêt. En l'absence d'énergie potentielle, éventuellement au sens généralisé (par exemple en présence d'un champ magnétique), la procédure ordinaire de quantification est stérile.

La même remarque vaut tout autant pour l'approche de Feynman, qui repose sur le formalisme lagrangien.

appeler symétrie *propre* ou *interne*, et qu'il convient de bien distinguer de l'invariance universelle résultant des hypothèses d'isotropie et d'homogénéité de l'espace. L'invariance euclidienne affirme que les propriétés de tout système sont indifférentes à la position ou à l'orientation du système, en conséquence de l'homogénéité et de l'isotropie de l'espace où il est plongé. Maintenant, pour un système donné, certains de ses éléments peuvent être indifférents à telle ou telle opération et c'est cette indifférence qui constitue la symétrie interne *propre*. Affirmer que les propriétés ne changent pas, c'est prendre en compte à la fois le principe euclidien et la symétrie interne propre : la seconde repose sur le premier. Pour une particule en l'absence de tout champ de force, la symétrie propre se fond avec la symétrie galiléenne.

Par exemple, au voisinage de la surface terrestre, comptant positivement vers le haut l'altitude z, le potentiel du champ de pesanteur[19] est $U(z) = gz$, l'énergie potentielle est $V(z) = mgz$. La fonction $V(z)$ a une propriété remarquable : elle ne dépend que de la variable verticale z et non des deux coordonnées horizontales x et y (physiquement : tout déplacement horizontal ne donne lieu à aucun travail de la force de pesanteur). Il en résulte que les propriétés d'équilibre ne dépendent que de z : l'équilibre dans un plan perpendiculaire à Oz est indifférent aux valeurs de x et de y. Ces deux coordonnées absentes de l'énergie potentielle sont appelées variables *cycliques*[20] et entraînent la conservation de l'impulsion dans un plan horizontal (p_x et p_y sont des constantes du mouvement). La symétrie correspondante est une symétrie de translation parallèlement au plan xOy : l'image de tout point d'équilibre dans une telle translation est aussi un point d'équilibre. Ceci constitue un exemple de propriété invariante, ici dans une opération géométrique : c'est l'une des symétries du système. La situation est identique pour une particule de charge q plongée dans un champ électrique extérieur $\vec{\mathcal{E}}$. Le Lagrangien s'écrit alors :

$$L = \frac{1}{2}\, m\vec{v}^{\,2} - (-q\vec{\mathcal{E}}.\vec{r}) \ . \tag{17.18}$$

Le champ définit une direction remarquable : il est tout indiqué, pour la commodité, de choisir l'un des axes du repère (cartésien) parallèle à ce champ, soit Oz pour fixer les idées. Il vient alors :

$$\boxed{L = \frac{1}{2}\, m\vec{v}^{\,2} + q\mathcal{E}z} \tag{17.19}$$

et on se retrouve exactement dans la même situation que pour la pesanteur, étudiée localement à la surface terrestre – ou plus généralement au voisinage de tout point dont la distance au centre de la Terre est donnée une fois pour toutes.

[19] C'est la linéarisation du terme $-G\frac{M}{d}$ avec $d = R_{\mathrm{T}} + z$ et $z \ll R_{\mathrm{T}}$, soit $-G\frac{M}{R_{\mathrm{T}}} + G\frac{M}{R_{\mathrm{T}}^2}z \equiv \mathrm{C^{ste}} + gz$. Dans les notations utilisées p. 663, on a $U = -G\frac{M}{R_{\mathrm{T}}} + G\frac{M}{R_{\mathrm{T}}^2}z \simeq \mathrm{C^{ste}} + gz$ et $V \simeq mU \equiv mgz$. R_{T} est le rayon de la Terre, G est la constante de gravitation universelle :

$$\boxed{R_{\mathrm{T}} \simeq 6\,400\,\mathrm{km}\ , \qquad G = 6,674\,28 \times 10^{-11}\,\mathrm{m^3kg^{-1}\,s^{-2}}\ , \qquad g \simeq 9,81\,\mathrm{m\,s^{-2}}} \tag{17.17}$$

[20] L'origine de cette terminologie tient au fait que, très souvent, les variables absentes de L (ou de H) sont des angles, définis à 2π près.

Autre exemple : l'atome d'hydrogène dans l'approximation du noyau infiniment massif. Le système physique d'étude se réduit à une seule particule (l'électron, de masse m et de charge $e < 0$) soumise à la force de Coulomb. Le Lagrangien est alors :

$$L = \frac{1}{2}m\vec{v}^2 - (-\frac{e'^2}{r}) \qquad (e'^2 = \frac{e^2}{4\pi\varepsilon_0}) \ . \tag{17.20}$$

Ici, $V(r) = -e'^2/r$ a la propriété remarquable de ne dépendre que de la distance électron – noyau, en aucune façon de la *direction* du rayon vecteur joignant les deux particules. Il en résulte une symétrie interne sphérique : toute rotation autour d'un axe passant par le noyau est une opération blanche, qui ne change en rien l'état du système (l'électron se moque de savoir ce que le physicien définit comme étant les trois axes d'un repère cartésien). Ici d'ailleurs, les coordonnées sphériques s'imposent, toujours pour des raisons de commodité ; il en faut trois : la distance au noyau r et les deux angles θ (latitude) et ϕ (azimut) définis comme d'habitude. Ces angles n'apparaissent pas dans le Lagrangien et sont de ce fait des variables cycliques. La symétrie, ici sphérique, produira une loi de conservation, celle du moment cinétique (*orbital*) de l'électron.

Supposons maintenant que l'on impose de l'extérieur un champ électrique, définissant toujours conventionnellement l'axe Oz. Le Lagrangien contient maintenant un terme de plus :

$$\boxed{L = \frac{1}{2}m\vec{v}^2 + \frac{e'^2}{r} + e\mathcal{E}z} \tag{17.21}$$

Visiblement, l'adjonction du champ réduit la symétrie interne : il n'y a maintenant invariance que par rapport aux rotations autour de l'axe Oz – et également par rapport aux réflexions par rapport à tout plan contenant ce même axe. On dit que le champ a *brisé* la symétrie sphérique, tout en préservant la symétrie cylindrique et la symétrie miroir ; au passage, on notera que le phénomène de brisure de symétrie[21] fonctionne par tout ou rien : un champ infinitésimal fait changer *qualitativement* la symétrie, ce qui se traduit par une réduction du nombre de variables cycliques. Avec les coordonnées cylindriques ρ, ϕ et z, qui sont de toute évidence appropriées à cette situation, on a :

$$\boxed{L = \frac{1}{2}m\vec{v}^2 + \frac{e'^2}{\sqrt{\rho^2+z^2}} + e\mathcal{E}z} \tag{17.22}$$

et il n'y a plus qu'une seule variable cyclique, l'angle azimutal ϕ. À cette variable absente de L sera associée la seule conservation de la projection le long de Oz du moment cinétique. Notons que, grâce à la symétrie de réflexion, les deux mouvements qui se correspondent dans cette opération sont physiquement équivalents ; pour deux tels mouvements, l'électron tourne dans un sens ou dans le sens contraire autour de Oz. Les propriétés physiques, en particulier la valeur de l'énergie, ne dépendront donc que du module de cette composante, pas de son signe[22].

[21] La notion de symétrie *brisée* se retrouve couramment à propos des systèmes non-linéaires (conduisant à celle de *bifurcation*), et dans la théorie des phénomènes critiques (par exemple : théorie de Landau).

[22] Il en va tout autrement avec un champ *magnétique* ; c'est pourquoi la levée de dégénérescence est

Dernier exemple, important : pour un système isolé, possédant par nature la symétrie de translation dans le temps, ni L ni H ne contiennent explicitement le temps. Alors, c'est l'énergie qui est une constante du mouvement (voir Tome I, chapitre 7).

En résumé, toute propriété de symétrie interne d'un système a pour contrepartie physique la conservation de certaines grandeurs dynamiques, et réciproquement. Techniquement, la symétrie se traduit par l'*absence* dans le Lagrangien de certaines coordonnées (ce sont précisément les *variables cycliques*), pourtant nécessaires par ailleurs pour fixer la position du système dans l'espace : cette absence produit automatiquement les lois de conservation des moments conjugués correspondants. En effet, dans les notations traditionnelles, les équations de Lagrange sont :

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad \Longleftrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}p_i = \frac{\partial L}{\partial q_i}} \tag{17.23}$$

$p_i \stackrel{\text{déf}}{=} \frac{\partial L}{\partial \dot{q}_i}$ étant le moment conjugué de q_i. Si une coordonnée q_{i_0} n'apparaît pas dans L, la dérivée $\frac{\partial L}{\partial q_{i_0}}$ est identiquement nulle ; par (17.23), il en résulte immédiatement que p_{i_0} est une constante du mouvement :

$$\boxed{q_{i_0} \text{ cyclique} \quad \Longleftrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}p_{i_0} = 0 \quad \Longleftrightarrow \quad p_{i_0} = \mathrm{C^{ste}}} \tag{17.24}$$

On doit à Emmy Noether[23] [11] d'avoir reconnu et formalisé le lien entre symétrie et constantes du mouvement, constituant le *Théorème de Noether.* D'une façon un peu formelle, ce théorème s'énonce ainsi : *À toute symétrie locale est associée la conservation d'un courant.* On sait en effet que toute loi de conservation peut s'exprimer de façon locale, l'équation $\partial_t\rho+\vec{\nabla}\vec{j} = 0$ pour la charge électrique étant l'exemple sans doute le plus familier (une telle équation est souvent appelée *équation de continuité* ; on peut préférer l'expression *équation de conservation*). Ce théorème vaut tout autant en classique qu'en quantique, et constitue le fondement de la démonstration de résultats majeurs. En théorie des champs par exemple, le théorème de Noether permet d'établir une relation importante, appelée *identité de Ward - Takahashi* ; celle-ci joue un rôle essentiel et a permis, notamment, de démontrer que le théorème de Goldstone est vrai à tous les ordres de la théorie des perturbations. Ce théorème est le fondement de toutes les théories de jauge[24], mais se manifeste également en Physique des basses énergies. Il stipule que toute brisure d'une symétrie continue s'accompagne de l'apparition d'excitations de basse énergie ; les magnons étudiés dans le chapitre 30, section 30.6, sont une conséquence, pour un ferromagnétique, de la brisure de symétrie de rotation.

seulement partielle pour l'effet Stark (champ électrique), alors qu'elle est totale pour l'effet Zeeman (champ magnétique). Physiquement la différence vient du fait que le couplage magnétique implique une *vitesse* (la brique élémentaire du magnétisme est une boucle de courant), qui change de signe quand on effectue la symétrie miroir. Le lien entre symétrie et dégénérescence est expliqué dans la section 17.6.

[23]Emmy Noether (1882-1935) était la fille du mathématicien Max Noether, qui apporta une contribution importante à la géométrie algébrique. Elle ne put faire ses études, notamment à Göttingen, que comme auditrice libre puisque les femmes à l'époque n'étaient pas acceptées comme étudiant(e)s. Grâce à l'intervention de Hilbert, elle put enseigner de façon officieuse, mais n'obtint jamais de poste. Mise en retraite d'office en 1933 par le régime nazi, elle se réfugia aux États-Unis et enseigna à Princeton.

[24]Pour une introduction très pédagogique, voir [12].

17.2 Opérateurs unitaires, opérateurs antiunitaires

Tous les opérateurs Ω considérés jusqu'ici ont été supposés linéaires, c'est-à-dire satisfaisant par définition :

$$\boxed{\Omega(\lambda|\psi\rangle + \mu|\phi\rangle) \overset{\text{déf}}{=} \lambda\Omega|\psi\rangle + \mu\Omega|\phi\rangle \qquad \forall\,\lambda,\,\mu\,\in\,\mathbb{C}} \tag{17.25}$$

en d'autres termes, un opérateur linéaire *commute* avec tous les scalaires du corps de l'espace vectoriel. L'adjoint $\Omega^\dagger$ d'un tel opérateur est défini par (voir Tome I, chapitre 13, éq. (13.11)) :

$$\boxed{(\Omega^\dagger|\psi\rangle,\,|\phi\rangle) \overset{\text{déf}}{=} (|\psi\rangle,\,\Omega|\phi\rangle)} \tag{17.26}$$

Un opérateur inversible U est dit *unitaire* ssi son adjoint $U^\dagger$ est égal à son inverse : $U^\dagger = U^{-1}$; un opérateur unitaire conserve le produit scalaire (voir Tome I, chapitre 13, éq. (13.59)). Rappelons que tout opérateur unitaire U peut s'écrire $U = \mathrm{e}^{\mathrm{i}A}$ où A est un opérateur hermitique (voir Tome I, problème 13.9.9) ; par ailleurs, on rencontrera souvent dans la suite la combinaison $UBU^\dagger$, où B est un certain opérateur, soit $\mathrm{e}^{\mathrm{i}A}B\mathrm{e}^{-\mathrm{i}A}$. On a vu que cette quantité peut se développer en puissances entières de A (voir problème 13.9.3, éq. (13.204)), développement rappelé ici pour la commodité du lecteur :

$$\mathrm{e}^{\mathrm{i}A}B\mathrm{e}^{-\mathrm{i}A} = B + \frac{\mathrm{i}}{1!}[A,\,B] + \frac{\mathrm{i}^2}{2!}[A,\,[A,\,B]] + \frac{\mathrm{i}^3}{3!}[A,\,[A,\,[A,\,B]]] + ... \tag{17.27}$$

Si l'un des commutateurs multiples est nul, le développement s'arrête à un rang fini ; c'est par exemple le cas avec $A = q$, $B = p$ puisque $[q,\,p] = \mathrm{i}\hbar\mathbf{1}$. On a ainsi :

$$\mathrm{e}^{\mathrm{i}kq}p\mathrm{e}^{-\mathrm{i}kq} = p + \frac{\mathrm{i}k}{1!}[q,\,p] + 0 = p - \hbar k\mathbf{1} \;; \tag{17.28}$$

le même phénomène se produit avec les opérateurs a et $a^\dagger$:

$$\mathrm{e}^{\mathrm{i}\alpha a}a^\dagger\mathrm{e}^{-\mathrm{i}\alpha a} = a^\dagger + \frac{\mathrm{i}\alpha}{1!}[a,\,a^\dagger] + 0 = a^\dagger + \mathrm{i}\alpha\mathbf{1} \;. \tag{17.29}$$

Toutes les symétries envisagées dans la suite conduiront à l'introduction d'opérateurs unitaires, à l'exception notable de celle liée au renversement du temps ; c'est pourquoi il convient de donner ici quelques définitions utiles pour la suite. De fait, il existe aussi des opérateurs *antilinéaires*, notés génériquement K, satisfaisant :

$$\boxed{K(\lambda|\psi\rangle + \mu|\phi\rangle) \overset{\text{déf}}{=} \lambda^*K|\psi\rangle + \mu^*K|\phi\rangle \qquad \forall\,\lambda,\,\mu\,\in\,\mathbb{C}} \tag{17.30}$$

de cette définition, il résulte immédiatement qu'un scalaire $\lambda \in \mathbb{C}$ ne commute pas avec un opérateur antilinéaire[25] :

$$K\,\lambda|\phi\rangle = \lambda^*K\,|\phi\rangle \quad \Longleftrightarrow \quad [K,\,\lambda] = (\lambda^* - \lambda)K \;; \tag{17.31}$$

[25] Il est tout à fait possible de définir la matrice M d'un opérateur antilinéaire sur une base donnée, selon la prescription habituelle : on porte successivement en colonne les transformés des vecteurs de la base. Toutefois, le vecteur transformé par l'opérateur d'un vecteur de composantes $\{c_n\}$ ne s'obtient pas simplement en multipliant la matrice M par la matrice colonne des c_n, car il faut en plus, conformément à (17.31), conjuguer ces derniers.

le commutateur n'est pas nul, sauf dans le cas où λ est réel.

L'adjoint d'un opérateur antilinéaire se définit comme suit :

$$\boxed{(K^\dagger|\psi\rangle,\,|\phi\rangle) \stackrel{\text{déf}}{=} (|\psi\rangle,\,K|\phi\rangle)^*} \tag{17.32}$$

Par rapport à la définition (17.26), on notera la conjugaison complexe au second membre. S'agissant d'opérateurs antilinéaires, la notation de Dirac doit être manipulée prudemment. En effet, en ajoutant des parenthèses dont la nécessité va apparaître clairement, le premier membre de (17.32) est :

$$(K^\dagger|\psi\rangle,\,|\phi\rangle) \equiv \langle K^\dagger\psi|\phi\rangle = (\langle\psi|K)|\phi\rangle\ . \tag{17.33}$$

D'un autre côté, suivant la définition (17.32), ce premier membre est égal à $[\langle\psi|(K|\phi\rangle]^*$; d'où, pour un opérateur antilinéaire :

$$\boxed{(\langle\psi|K)|\phi\rangle = \langle\psi|(K|\phi\rangle)^*} \tag{17.34}$$

où il saute aux yeux que les parenthèses ne peuvent être omises[26]. Au risque de la redondance, cela vaut la peine de récrire la même suite d'égalités pour un opérateur linéaire ; la seule différence étant l'absence de conjugaison dans la définition de l'adjoint, on a d'abord :

$$(\Omega^\dagger|\psi\rangle,\,|\phi\rangle) \equiv \langle\Omega^\dagger\psi|\phi\rangle) = (\langle\psi|\Omega)|\psi\rangle\ ; \tag{17.35}$$

suivant (17.26), le premier membre est égal à $\langle\psi|(\Omega|\phi\rangle)$, d'où $(\langle\psi|\Omega)|\psi\rangle = \langle\psi|(\Omega|\phi\rangle)$, confirmant que, pour un opérateur linéaire, les parenthèses ne servent à rien et peuvent être omises – comme on l'a toujours fait.

Par définition, un opérateur antiunitaire est un opérateur antilinéaire inversible dont l'adjoint est égal à l'inverse :

$$\boxed{\mathcal{U}\text{ est antiunitaire} \quad\Longleftrightarrow\quad \mathcal{U}\text{ est antilinéaire et } \mathcal{U}^\dagger = \mathcal{U}^{-1}} \tag{17.36}$$

Pour un opérateur $\mathcal{U}$ antiunitaire, la définition (17.32) donne :

$$(\mathcal{U}^{-1}|\psi\rangle,\,|\phi\rangle) = (|\psi\rangle,\,\mathcal{U}|\phi\rangle)^*\ ; \tag{17.37}$$

en posant $|\psi'\rangle \stackrel{\text{déf}}{=} \mathcal{U}^{-1}|\psi\rangle$, il vient $(|\psi'\rangle,\,|\phi\rangle) = (\mathcal{U}|\psi'\rangle,\mathcal{U}|\phi\rangle)^*$. Ainsi, un opérateur antiunitaire a la propriété de transformer le produit scalaire en son complexe conjugué :

$$\boxed{\mathcal{U}\text{ antiunitaire} \quad\Longleftrightarrow\quad (\mathcal{U}|\psi\rangle,\mathcal{U}|\phi\rangle) = (|\psi\rangle,\,|\phi\rangle)^* \qquad \forall\,|\psi\rangle,\,|\phi\rangle} \tag{17.38}$$

On rencontrera un opérateur antiunitaire lors de l'étude de l'invariance par renversement du temps (section 17.5).

[26] À tout prendre, la notation du produit scalaire sous forme d'un couple ordonné, $(|\flat\rangle,\,|\sharp\rangle)$, n'est pas plus compliquée à écrire que la notation de Dirac – en cas de doute, mieux vaut y revenir.

17.3 Symétrie spatiale continue : translations et rotations

Une symétrie spatiale *continue* se caractérise par le fait que l'on peut lui associer des transformations *infinitésimales*, arbitrairement proches de la transformation identité : on peut translater un point sur une longueur aussi petite que l'on veut, ou le faire tourner d'un angle arbitrairement petit. S'en tenant strictement au plan géométrique[27], deux types de transformations jouent un rôle de tout premier plan en Mécanique (classique ou quantique)[28], les translations et les rotations. À ces symétries continues s'opposent les symétries discrètes (la réflexion, la parité, le renversement du temps,...), pour lesquelles on ne peut définir d'opérations infinitésimales[29]. Les translations et les rotations appartiennent à la classe des *déplacements*, opérations géométriques qui ne changent ni les longueurs, ni les angles, ni la chiralité. Il existe aussi, dans le plan $\mathbb{R}^2$, des transformations *conformes* qui modifient les longueurs mais préservent les angles[30].

Tout naturellement dans la suite, on se place dans la représentation-q, puisqu'il sera essentiellement question de transformations accomplies au départ dans l'espace physique (en général $\mathbb{R}^3$). Toutefois, l'expression formelle des opérateurs $\mathcal{R}$ agissant dans l'espace des états associés aux opérations R effectuées dans $\mathbb{R}^3$ est la même, quelle que soit la représentation finalement retenue dans un contexte où celle-ci se révèle plus appropriée ; par exemple, si l'opérateur de translation $\mathcal{T}_{\vec{a}}$ est effectivement l'opérateur différentiel $\mathrm{e}^{-\vec{a}.\vec{\nabla}}$ en représentation-q, et le simple facteur multiplicatif $\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\vec{p}.\vec{a}}$ en représentation-p, $\mathcal{T}_{\vec{a}}$ est dans tous les cas égal à l'opérateur $\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\vec{p}.\vec{a}}$ (voir (17.68)). On aura reconnu dans ces relations les correspondances qui existent entre une fonction et ses dérivées selon que l'on se place dans l'espace direct ou dans l'espace réciproque introduit par la transformation de Fourier : la différentiation dans l'espace direct a pour image la simple multiplication dans l'espace réciproque[31]. En définitive, tous les résultats obtenus en termes d'opérateurs ont une validité générale, une représentation ou une autre étant adoptée en temps utile pour la commodité des calculs, tout comme en géométrie élémentaire le choix d'un repère adapté rend les calculs plus simples.

17.3.1 Translations

Un système étant donné, la symétrie d'invariance par translation dans l'espace se traduit par le fait que certaines coordonnées rectangulaires (x, y ou z) ne figurent pas dans l'expression du Lagrangien et donc du Hamiltonien. Afin de fixer les idées, le système

[27]On verra qu'il existe des transformations de jauge infinitésimales.

[28]Assez récemment, une autre symétrie a révélé son importance : la similitude, ou symétrie affine (dilatation ou contraction des longueurs). Cette symétrie donne lieu à l'invariance d'échelle sur laquelle est construite la zoologie des structures fractales.

[29]En outre, la réflexion ne respecte pas la chiralité, au contraire de toute rotation, même discrète.

[30]Les transformations conformes sont associées aux fonctions *holomorphes* [13] ; elles jouent un rôle important pour la Physique à deux dimensions dans des domaines très variés (effets de bord dans un condensateur plan, effets de taille finie dans la théorie des transitions de phase, théorie des champs, ...).

[31]À des détails près, ceci vaut aussi pour la transformation de Laplace.

de référence pour l'exemple sera une particule dans un champ constant (gravité, champ électrique,...), possédant l'énergie potentielle $V(z) = -Kz$. Le Lagrangien est :

$$L = \frac{1}{2} m\vec{v}^2 + Kz \ . \tag{17.39}$$

Les moments conjugués $p_u = \partial L/\partial \dot{q}_u$, ($u = x, y,z$) sont égaux à mv_u. Les équations de Lagrange donnent :

$$\frac{\mathrm{d}}{\mathrm{d}t}(mv_u) = \frac{\partial L}{\partial u} = -\frac{\partial V}{\partial u} \equiv F_u \ , \tag{17.40}$$

et s'identifient, par construction du Lagrangien, à l'équation fondamentale de la dynamique projetée sur les trois axes. Comme on l'a vu, l'absence de x et de y dans L conduit immédiatement à la conservation de p_x et p_y :

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}(mv_x) = 0 \qquad \frac{\mathrm{d}}{\mathrm{d}t}(mv_y) = 0} \tag{17.41}$$

Le caractère cyclique des deux coordonnées x et y assure *ipso facto* la conservation des projections correspondantes de la quantité de mouvement.

Évidemment, tout ceci se retrouve dans la formulation hamiltonienne. Le Hamiltonien se construit comme d'habitude ; après avoir trouvé les moments conjugués en utilisant leur définition $p_i \stackrel{\text{déf}}{=} \partial L/\partial \dot{q}_i$, on fabrique la combinaison[32] :

$$H = \sum_i p_i \dot{q}_i - L \tag{17.42}$$

que l'on exprime à l'aide des p_i et des q_i. On trouve ainsi facilement :

$$\boxed{H = \frac{\vec{p}^2}{2m} + V(z)} \tag{17.43}$$

Le Hamiltonien H est la somme de l'énergie cinétique et de l'énergie potentielle : c'est l'énergie totale du système dans le champ considéré, indépendante du temps pour un système isolé (conservation de l'énergie se traduisant par l'existence de l'intégrale première $H(q_i, p_i) = E$ le long de la trajectoire, constante dont la valeur est fixée une fois pour toutes par les conditions initiales prescrites). La conservation de p_x et de p_y découle immédiatement du fait que L et H ne dépendent ni de x ni de y ; soit par (17.41), soit à partir des équations de Hamilton :

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \ , \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \ ; \tag{17.44}$$

on déduit alors de (17.43) que p_x et p_y sont des constantes le long de la trajectoire réellement suivie.

[32]C'est une transformation de Legendre, voir Tome I, chapitre 7.

La notion de variables cycliques se transpose en Mécanique quantique. En effet, il suffit de revenir aux équations de Heisenberg (Tome I, chapitre 14, sous-section 14.1.2) :

$$\boxed{\mathrm{i}\hbar\,\frac{\mathrm{d}A_{\mathrm{H}}}{\mathrm{d}t} = [A_{\mathrm{H}}, H] \equiv [A, H]_{\mathrm{H}}} \tag{17.45}$$

qui sont l'une des illustrations de la correspondance entre le crochet de Poisson $\{A,\, B\}$ et $\frac{1}{\mathrm{i}\hbar}[A,\, B]$, puisque l'équation d'évolution classique de l'"observable" A (supposée indépendante du temps) est $\frac{\mathrm{d}A}{\mathrm{d}t} = \{A,\, H\}$ (voir Tome I, éq. (7.89)). Pour une particule dans $\mathbb{R}^3$, et choisissant pour A les coordonnées et les moments conjugués, on obtient ainsi les six équations :

$$\mathrm{i}\hbar\,\frac{\mathrm{d}u_{\mathrm{H}}}{\mathrm{d}t} = [u,\, \frac{p_u^2}{2m}]_{\mathrm{H}}\ , \qquad \mathrm{i}\hbar\,\frac{\mathrm{d}p_{u,\,\mathrm{H}}}{\mathrm{d}t} = [p_u,\, V(\vec{r})]_{\mathrm{H}} \qquad (u = x, y, z) \tag{17.46}$$

qui, après calcul, donnent :

$$\boxed{\dot{u}_{\mathrm{H}} = \frac{1}{m}\,p_{u,\,\mathrm{H}} \qquad \dot{p}_{u,\,\mathrm{H}} = -\left(\frac{\partial V}{\partial u}\right)_{\mathrm{H}} \qquad (u = x, y, z)} \tag{17.47}$$

Ces équations du mouvement quantique sont en tout point analogues formellement à leurs homologues classiques en raison de la correspondance rappelée ci-dessus mais, bien sûr, impliquent des opérateurs. Supposons maintenant que certaines variables sont absentes de H ; dans l'exemple ci-dessus, mais maintenant traité quantiquement, donc décrit par le Hamiltonien (17.43) où les variables classiques ont été remplacées par leurs opérateurs, x et y ne figurent pas dans H. De ce simple fait, les commutateurs $[x,\, H]$ et $[y,\, H]$ sont nuls, d'où $\dot{p}_{x,\,\mathrm{H}} = \dot{p}_{y,\,\mathrm{H}} = 0$: $p_{x,\mathrm{H}}$ et $p_{y,\mathrm{H}}$ sont des constantes du mouvement au sens quantique ; les opérateurs $p_{x,\,\mathrm{H}}(t)$ et $p_{y,\,\mathrm{H}}(t)$ coïncidant avec p_x et p_y, les valeurs moyennes des observables correspondantes sont des constantes dans le temps, quel que soit l'état de départ (voir Tome I, éqs. (14.67) et (14.68)).

Bien évidemment, l'invariance euclidienne a des conséquences sur le formalisme quantique, qui vont se manifester de façon plus abstraite : toute prévision quantique doit satisfaire le principe euclidien, ce qui impose un mode de transformation bien déterminé pour la fonction d'onde et les observables. En toute généralité, l'invariance euclidienne contraint très fortement la structure algébrique de la Mécanique quantique (on pourrait même parler de structure *géométrique*, tant elle est liée à la géométrie au sens ordinaire). De surcroît, toute symétrie interne se traduit sur les attributs quantiques d'un système donné ; pour l'exemple cité plus haut, on va voir que le Hamiltonien lui-même est invariant dans certaines opérations définies dans l'espace des états, images abstraites des translations géométriques parallèles au plan $x\mathrm{O}y$ effectuées dans l'espace physique $\mathbb{R}^3$.

Le système quantique décrit par (17.43) possède intrinsèquement une symétrie d'invariance par translation perpendiculairement au champ. Ceci signifie que toute translation du système physique[33] n'affecte en rien les prévisions effectuées à son propos :

[33]Lorsqu'il s'agit d'étudier l'invariance par symétrie, il y a toujours deux points de vue possibles :

- soit on effectue un changement de repère, sans toucher au système, le second repère se déduisant

comme l'état du système ne change pas dans une telle translation, celui-ci se moque du choix du repère qui est fait pour la commodité de calcul ; trouver des résultats différents avant et après translation serait une absurdité puisque cela reviendrait à dire qu'un système et le même système ont des propriétés différentes.

Considérons d'abord une translation *quelconque* du système, $T_{\vec{a}}$, caractérisée par le vecteur $\vec{a}$. Dans cette hypothèse, la propriété d'invariance impliquée est uniquement celle qui découle du principe euclidien – au contraire, quand $\vec{a}$ est perpendiculaire à l'axe Oz, c'est alors la symétrie propre du système qui est en jeu, toujours assise sur l'invariance galiléenne universelle. Dans une translation quelconque, un vecteur $\vec{r}$ se transforme en un vecteur $\vec{r}\,'$:

$$\boxed{\vec{r}\,' \stackrel{\text{déf}}{=} T_{\vec{a}}\,\vec{r} = \vec{r} + \vec{a}} \tag{17.48}$$

Avant translation, la fonction d'onde est $\Psi(\vec{r},\,t)$; après translation, c'est *a priori* une nouvelle fonction, $\Psi'(\vec{r},\,t)$, le paramètre temps étant bien sûr le même. Ces deux fonctions d'onde n'ont évidemment aucune raison de coïncider par elles-mêmes : en revanche, ce qui importe c'est l'identité des prévisions faites avec l'une et avec l'autre. Si elles ne coïncident pas, il doit toutefois exister un opérateur agissant dans l'espace des fonctions et transformant l'une en l'autre. Ceci étant admis, on peut écrire formellement, en termes de vecteurs (kets), pour deux états quelconques $|\Psi\rangle$ et $|\Phi\rangle$:

$$|\Psi'\rangle \stackrel{\text{déf}}{=} \mathcal{T}_{\vec{a}}\,|\Psi\rangle \ , \qquad |\Phi'\rangle \stackrel{\text{déf}}{=} \mathcal{T}_{\vec{a}}\,|\Phi\rangle \ , \tag{17.49}$$

$\mathcal{T}_{\vec{a}}$ étant l'opérateur agissant dans l'espace des états $\mathcal{E}$ et représentant la translation $T_{\vec{a}}$ effectuée dans l'espace physique ; ces deux opérations sont de nature bien différente, ce que traduit le symbole les désignant. Dans le langage des fonctions d'onde, les relations de définition (17.49) se transcrivent comme suit :

$$\Psi'(\vec{r},\,t) \stackrel{\text{déf}}{=} \mathcal{T}_{\vec{a}}\,\Psi(\vec{r},\,t) \ , \qquad \Phi'(\vec{r},\,t) \stackrel{\text{déf}}{=} \mathcal{T}_{\vec{a}}\,\Phi(\vec{r},\,t) \ . \tag{17.50}$$

En Mécanique quantique, toute prévision s'énonce en terme de probabilités, lesquelles s'expriment d'une façon générale par le module au carré du produit scalaire approprié. L'invariance requise par la symétrie euclidienne de translation impose que l'on ait, $\forall\, \vec{a}$:

$$|(|\Psi'\rangle,\,|\Phi'\rangle)|^2 = |(|\Psi\rangle,\,|\Phi\rangle)|^2 \quad \Longleftrightarrow \quad |(|\Psi'\rangle,\,|\Phi'\rangle)| = |(|\Psi\rangle,\,|\Phi\rangle)| \tag{17.51}$$

et rien de plus : l'invariance ne demande pas l'égalité des produits scalaires $(|\Psi'\rangle,\,|\Phi'\rangle)$ et $(|\Psi\rangle,\,|\Phi\rangle)$, auquel cas l'opérateur $\mathcal{T}_{\vec{a}}$ serait forcément unitaire, mais seulement, d'après (17.51), l'égalité de leurs *modules*. En définitive, l'invariance requise est du type *projectif* : tous les *rayons* $\mathrm{e}^{\mathrm{i}\alpha}\,|\Psi\rangle$ ont *a priori* exactement le même contenu physique.

du premier par une transformation géométrique (point de vue "passif"),

- soit on effectue la transformation inverse sur le système lui-même (point de vue "actif").

S'agissant de rotations, par exemple, tourner le repère dans un sens ou tourner le système dans l'autre sens revient au même, précisément grâce au principe de relativité. Dans la suite, sauf mention du contraire, on adopte le point de vue actif : c'est le système que l'on translate, que l'on fait tourner, etc.

Il est possible de montrer (théorème de Wigner[34] ([14], [4] p. 372, [15] p. 540, théorème III)) que la condition (17.51) impose en général à tout opérateur représentant une opération de symétrie d'être soit unitaire (c'est de loin le cas le plus répandu) – tout produit scalaire est alors invariant –, soit antiunitaire – auquel cas tout produit scalaire est transformé en son complexe conjugué. Pour une symétrie continue, l'opérateur est forcément unitaire, puisque l'on doit pouvoir tendre par continuité vers la transformation identité, trivialement unitaire[35]. Ce théorème peut d'ailleurs se démontrer au cas par cas en utilisant des arguments physiques, comme on va le voir.

Comme exemple de prévision, considérons la probabilité d'observer la particule au voisinage du point $\vec{r}$; la densité de probabilité correspondante est donnée par le module au carré de la fonction d'onde. Une fois que l'on a translaté le système, c'est donc $|\Psi'(\vec{r},\, t)|^2$ qui donne la densité de probabilité de présence au point $\vec{r}$; l'invariance (euclidienne) exige que ce nombre soit le même que celui calculé, pour le système avant translation, au point qui est l'antécédent de $\vec{r}$ dans la translation. Il faut donc[36] :

$$|\Psi'(\vec{r},\, t)|^2 \equiv |\mathcal{T}_{\vec{a}}\, \Psi(\vec{r},\, t)|^2 = |\Psi(T_{\vec{a}}^{-1}\, \vec{r},\, t)|^2 \ , \tag{17.53}$$

égalité qui entraîne :

$$\Psi'(\vec{r},\, t) = \mathrm{e}^{\mathrm{i}\alpha}\, \Psi(T_{\vec{a}}^{-1}\, \vec{r},\, t) \tag{17.54}$$

où α est une phase dépendant *a priori* de l'espace.

En fait, parce que la translation dans l'espace n'implique pas la vitesse (donc pas le moment conjugué), la phase α est en réalité une constante. En effet, avec la relation (17.54), la valeur moyenne du moment conjugué dans l'état transformé est $\langle \vec{p} \rangle' \stackrel{\text{déf}}{=} \langle \Psi' | \vec{p} | \Psi' \rangle$, soit :

$$\langle \vec{p} \rangle' = \int_{\mathbb{R}^3} \mathrm{e}^{-\mathrm{i}\alpha}\, \Psi^*(T_{\vec{a}}^{-1}\, \vec{r},\, t)\, \vec{p}\, \mathrm{e}^{+\mathrm{i}\alpha}\, \Psi(T_{\vec{a}}^{-1}\, \vec{r},\, t)\, \mathrm{d}^3 r \ . \tag{17.55}$$

---

[34]Eugene Paul Wigner (1902-1995), né à Budapest, était à la fois physicien et mathématicien. Il fut lauréat Nobel en 1963 *"for his contributions to the theory of the atomic nucleus and the elementary particles, particularly through the discovery and application of fundamental symmetry principles"*. On lui doit de nombreux théorèmes importants sur les symétries en Mécanique quantique (autre exemple : le théorème de Wigner - Eckart, voir chapitre 18, section 18.6), ainsi qu'une théorie statistique de la distribution des niveaux d'énergie des noyaux, utilisant sa théorie des matrices aléatoires gaussiennes développée avec F.J. Dyson.

Il est également connu pour avoir publié un article au titre quelque peu provocateur [16].

[35]L'antiunitarité ne peut être associée qu'à une opération ne pouvant *continûment* se réduire à ne rien faire du tout (caractère "discret", "quantifié", de l'opération considérée). Le renversement du temps en est un exemple : renverser ou ne pas renverser le temps est visiblement une opération par tout ou rien, tout comme une symétrie-miroir. Toutefois, ne pas en déduire que toute symétrie discrète est associée à un opérateur antiunitaire : la parité ou la symétrie-miroir sont représentées par des opérateurs unitaires.

[36]L'équation (17.53) est bien la généralisation formelle de la transformée d'une fonction. En termes plus simples, soit une fonction d'une seule coordonnée, $f(x)$, représentant un certain attribut d'un système. Si le système est translaté de x_0, $x' \equiv T_{x_0}\, x = x + x_0$ et le système est alors décrit par la fonction $f(x - x_0)$ (on prend en bloc $f(x)$ et on la déplace de x_0). On a bien :

$$f'(x) \equiv \mathcal{T}_{x_0}\, f(x) = f(x - x_0) = f(T_{x_0}^{-1}\, x) \ . \tag{17.52}$$

Le facteur de phase $\mathrm{e}^{-\mathrm{i}\alpha}$ est au plus une fonction de $\vec{r}$; le faisant passer à droite de $\Psi^*(T_{\vec{a}}^{-1}\,\vec{r},\,t)$, on fait apparaître la combinaison $\mathrm{e}^{-\mathrm{i}\alpha}\,\vec{p}\,\mathrm{e}^{+\mathrm{i}\alpha}$, qui vaut[37] :

$$\mathrm{e}^{-\mathrm{i}\alpha}\,\vec{p}\,\mathrm{e}^{+\mathrm{i}\alpha} = \vec{p} - \mathrm{i}[\alpha,\vec{p}] + \frac{\mathrm{i}^2}{2}\,[\alpha,[\alpha,\vec{p}]] + \ldots = \vec{p} - \hbar\,[\alpha,\vec{\nabla}] = \vec{p} + \hbar\vec{\nabla}\alpha \ , \qquad (17.56)$$

le report de cette expression dans (17.55) donne :

$$\langle\vec{p}\rangle' = \langle\vec{p}\rangle + \hbar\,\langle\vec{\nabla}\alpha\rangle \ . \qquad (17.57)$$

Ainsi, tant que l'on maintient la phase α variable dans l'espace, on trouve cette absurdité : translater un système dans l'espace modifie la valeur moyenne de sa vitesse. La seule issue est donc $\alpha = \mathrm{C}^{\mathrm{ste}}$; α étant une constante, on peut toujours s'en affranchir et prendre celle-ci égale à zéro. En conséquence, l'égalité (17.54) se simplifie, la fonction d'onde transformée étant donnée par :

$$\boxed{\Psi'(\vec{r},\,t) \stackrel{\text{déf}}{=} \mathcal{T}_{\vec{a}}\Psi(\vec{r},\,t) = \Psi(T_{\vec{a}}^{-1}\,\vec{r},\,t)} \qquad (17.58)$$

Cela étant, on voit immédiatement que l'opérateur $\mathcal{T}_{\vec{a}}$ est unitaire, comme il se doit puisqu'il existe des translations infinitésimales. En effet, pour deux fonctions quelconques Ψ et Φ, on a maintenant :

$$\int_{\mathbb{R}^3} \Psi'^*(\vec{r},\,t)\,\Phi'(\vec{r},\,t)\,\mathrm{d}^3r = \int_{\mathbb{R}^3} \Psi^*(T_{\vec{a}}^{-1}\,\vec{r},\,t)\,\Phi(T_{\vec{a}}^{-1}\,\vec{r},\,t)\,\mathrm{d}^3r \ ; \qquad (17.59)$$

dans le second membre, il suffit de faire un changement de variable d'intégration en posant $\vec{\rho} = T_{\vec{a}}^{-1}\,\vec{r} = \vec{r} - \vec{a}$ pour arriver à[38] :

$$\int_{\mathbb{R}^3} \Psi'^*(\vec{r},\,t)\,\Phi'(\vec{r},\,t)\,\mathrm{d}^3r = \int_{\mathbb{R}^3} \Psi^*(\vec{\rho},\,t)\,\Phi(\vec{\rho},\,t)\,\mathrm{d}^3\rho \ , \qquad (17.60)$$

égalité montrant que le produit scalaire des états transformés $(|\Psi'\rangle,\,|\Phi'\rangle) \equiv \langle\Psi'|\Phi'\rangle$ est égal au produit scalaire des états avant transformation ; l'invariance du produit scalaire est la signature d'une transformation unitaire, par définition de celle-ci.

Des résultats précédents, on déduit la valeur moyenne de la coordonnée dans l'état translaté, notée[39] $\langle\vec{r}\rangle'$; par définition de la valeur moyenne :

$$\langle\vec{r}\rangle' = \int_{\mathbb{R}^3} \Psi'^*(\vec{r},\,t)\,\vec{r}\,\Psi'(\vec{r},\,t)\,\mathrm{d}^3r \ . \qquad (17.61)$$

D'après (17.48) et (17.58), il vient :

$$\langle\vec{r}\rangle' = \int_{\mathbb{R}^3} \Psi^*(\vec{\rho},\,t)\,(\vec{\rho} + \vec{a})\,\Psi(\vec{\rho},\,t)\,\mathrm{d}^3\rho \ , \qquad (17.62)$$

[37] voir Tome I, problème 13.9.3, éq.(13.204), ou le rappel (17.27) et sa conséquence immédiate (17.28).
[38] Noter que l'on utilise ici l'hypothèse suivant laquelle l'espace est infini dans toutes les directions.
[39] Bien remarquer que le $'$ est à *l'extérieur* du signe $\langle\ldots\rangle$ représentant la moyenne.

soit finalement (compte tenu de la normalisation de la fonction d'onde) :

$$\boxed{\langle\vec{r}\rangle' = \langle\vec{r}\rangle + \vec{a}} \tag{17.63}$$

où $\langle\vec{r}\rangle$ désigne la valeur moyenne dans l'état avant translation. Ce résultat, trivial à la réflexion, est bien ce que l'on attend. En ce qui concerne l'impulsion – représentée par un opérateur de dérivation d'espace, insensible à une translation donnée – on a bien sûr $\langle\vec{p}\rangle' = \langle\vec{p}\rangle$ comme il se doit.

Il est maintenant aisé d'écrire explicitement l'expression de l'opérateur de translation $\mathcal{T}_{\vec{a}}$. Pour cela, commençons par considérer une translation infinitésimale d'amplitude $\delta\vec{a}$; par définition de l'opérateur de translation (17.50) et d'après (17.58), on a :

$$\Psi'(\vec{r}) \equiv \mathcal{T}_{\delta\vec{a}}\Psi(\vec{r}) = \Psi(T_{\delta\vec{a}}^{-1}\,\vec{r}) = \Psi(\vec{r} - \delta\vec{a}) \ . \tag{17.64}$$

En développant au premier ordre, on obtient :

$$\forall\ \Psi :\ \mathcal{T}_{\delta\vec{a}}\Psi(\vec{r}) = \Psi(\vec{r}) \ - \ \delta\vec{a}.\vec{\nabla}\Psi(\vec{r}) + \ldots \equiv (\mathbf{1} + \frac{\vec{p}}{\mathrm{i}\hbar}.\delta\vec{a} + \ldots)\,\Psi(\vec{r}) \ . \tag{17.65}$$

Ceci montre que l'opérateur $\mathcal{T}_{\delta\vec{a}}$ est :

$$\boxed{\mathcal{T}_{\delta\vec{a}} = \mathbf{1} + \frac{\vec{p}}{\mathrm{i}\hbar}\,\delta\vec{a} + \mathcal{O}(\delta\vec{a}^{\,2})} \tag{17.66}$$

En raison de ce résultat, on dit que $\vec{p}$ est le *générateur* des translations. Pour avoir l'opérateur d'une translation finie, il suffit maintenant de multiplier une suite de translations, toutes parallèles entre elles en posant $\delta\vec{a} = \frac{1}{N}\,\vec{a}$ et en écrivant :

$$\mathcal{T}_{\vec{a}} = (\mathcal{T}_{\delta\vec{a}})^N = \Big(\mathbf{1} + \frac{\vec{p}}{\mathrm{i}\hbar}\,\frac{\vec{a}}{N} + \ldots\Big)^N \ . \tag{17.67}$$

Dans la limite $N \to \infty$, compte tenu de $\lim_{N\to\infty}[1+(x/N)+\mathcal{O}(N^{-2})]^N = \mathrm{e}^x$ on obtient :

$$\boxed{\mathcal{T}_{\vec{a}} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\vec{p}.\vec{a}}} \tag{17.68}$$

Visiblement, et comme il se doit, cet opérateur est unitaire puisque $\vec{p}$ est hermitique[40] et que $\vec{a}$ est un vecteur réel. À titre de vérification, calculons effectivement, en représentation-q, l'action de l'opérateur $\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}p_x a}$ sur une fonction quelconque $f(x)$ supposée analytique[41]. On a :

$$\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}p_x a}f(x) = \mathrm{e}^{-a\frac{\mathrm{d}}{\mathrm{d}x}}f(x) = \sum_{n\in\mathbb{N}}\frac{(-1)^n}{n!}(a\frac{\mathrm{d}}{\mathrm{d}x})^n\,f(x) \equiv \sum_{n\in\mathbb{N}}\frac{1}{n!}(-a)^n f^{(n)}(x) \ ; \tag{17.69}$$

[40] Rappelons que l'exactitude de cette affirmation repose sur la satisfaction implicite de *conditions aux limites* pour les fonctions sur lesquelles agissent l'opérateur en question (voir Tome I, note 35 p. 334). D'autres exemples d'une telle nécessité seront rencontrés dans la suite, voir notamment les discussions pp. 781 et 728, et le problème 18.7.2.

[41] Une fonction $f(z)$ est dite *analytique* en z_0 s'il existe un voisinage de z_0, $\mathcal{V}(z_0)$, dans lequel cette fonction admet un développement en série entière de la variable $(z - z_0)$:

$$\forall\, z \in \mathcal{V}(z_0),\ f(z) = \sum_{n\in\mathbb{N}} c_n(z - z_0)^n$$

on reconnaît à droite le développement de Taylor de $f(x-a)$, d'où $\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}p_x a}f(x) = f(x-a)$, en conformité avec la définition (17.58), qui s'écrit ici $\mathcal{T}_a f(x) = f(T_a^{-1}\,x)$; d'après (17.48), ceci est bien égal à $f(x-a)$.

De toute évidence, l'ensemble des translations peut être doté d'une structure de groupe par la loi de composition interne $T_{\vec{a}_1}T_{\vec{a}_2} \stackrel{\text{déf}}{=} T_{\vec{a}_1+\vec{a}_2}$, tout comme l'ensemble des opérateurs $\mathcal{T}_{\vec{a}}$ correspondants. La commutativité de la somme vectorielle $\vec{a}_1+\vec{a}_2$ montre que $T_{\vec{a}_1}T_{\vec{a}_2} = T_{\vec{a}_2}T_{\vec{a}_1}$: le groupe des translations est donc abélien. Cette propriété s'induit d'elle-même dans le groupe des opérateurs $\mathcal{T}_{\vec{a}}$, qui doit refléter *fidèlement* les propriétés des transformations dans l'espace. L'expression (17.68) traduit bien la propriété abélienne puisque les différentes composantes de $\vec{p}$ commutent entre elles[42] ; de ce fait, l'exponentielle de la somme est le produit des exponentielles et, dans ce produit, l'ordre des facteurs est sans importance :

$$\mathcal{T}_{\vec{a}} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\sum_{u=x,y,z}p_u a_u} = \prod_{u=x,y,z}\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}p_u a_u} \equiv \prod_{u=x,y,z}\mathcal{T}_{a_u}\ . \tag{17.70}$$

Tout comme le vecteur d'état change quand on translate le système, le Hamiltonien H change lui aussi, et se transforme en H'. La recette définissant le mode de calcul des valeurs moyennes étant définie sans référence à un repère quelconque, l'invariance euclidienne impose l'égalité des moyennes calculées d'une façon et d'une autre :

$$(|\Psi'\rangle,\,H'|\Psi'\rangle) = (|\Psi\rangle,\,H|\Psi\rangle) \quad\Longleftrightarrow\quad \langle\Psi|\mathcal{T}_{\vec{a}}^{\dagger}H'\mathcal{T}_{\vec{a}}|\Psi\rangle = \langle\Psi|H|\Psi\rangle\ . \tag{17.71}$$

Ceci étant vrai $\forall\ |\Psi\rangle$ (par exemple pour l'ensemble des vecteurs propres de H, qui forment une base complète), l'égalité s'élève au niveau des opérateurs, $H = \mathcal{T}_{\vec{a}}^{\dagger}H'\mathcal{T}_{\vec{a}}$ quel que soit $\vec{a}$. D'où :

$$\boxed{H' = \mathcal{T}_{\vec{a}}H\mathcal{T}_{\vec{a}}^{\dagger} = \mathcal{T}_{\vec{a}}H\mathcal{T}_{\vec{a}}^{-1} \qquad \text{(invariance euclidienne)}} \tag{17.72}$$

Le vecteur de translation étant quelconque, ceci est l'expression analytique du principe euclidien[43]. On retiendra l'idée suivante : dans une transformation le vecteur d'état change ; puisque toutes les prévisions doivent être invariantes au sens euclidien, il faut bien que les observables changent simultanément, de façon à compenser la variation du vecteur d'état.

Revenons maintenant au cas précis considéré au début, servant d'exemple à la symétrie (interne, propre) de translation parallèlement à Ox et à Oy ; le Hamiltonien H défini en (17.43) est invariant dans toute translation perpendiculaire à Oz, puisque celle-ci ne change que x et y, les deux coordonnées dont H ne dépend pas :

$$\forall\ \vec{a}\perp \mathrm{O}z\ :\quad H' = H \quad\Longleftrightarrow\quad \mathcal{T}_{\vec{a}\perp\mathrm{O}z}^{\dagger}H\mathcal{T}_{\vec{a}\perp\mathrm{O}z} = H\ . \tag{17.73}$$

[42]En représentation-q, $p_u = -\mathrm{i}\hbar\partial_u$, où $u = x,\,y,\,z$ est une coordonnée rectangulaire. On suppose partout que les fonctions sur lesquelles agissent les opérateurs différentiels p_u, ont les propriétés analytiques requises (théorème de Clairault, voir par exemple [17], p. 407) pour que toutes les dérivées croisées soient égales, assurant que $[p_u,\,p_v] = 0$ quels que soient u et v.

[43]Par exemple, soit un oscillateur harmonique dont le point d'équilibre est pris à l'origine ; le terme d'énergie potentielle dans H est donc $(m\omega^2/2)\,x^2$; si on translate le système de x_0, le point d'équilibre a maintenant l'abscisse x_0 et l'énergie potentielle figurant dans H' est $(m\omega^2/2)\,(x-x_0)^2$. On vérifie que l'équation (17.72) est bien satisfaite avec l'opérateur $\mathcal{T}_{x_0} = \mathrm{e}^{-x_0\,(\mathrm{d}/\mathrm{d}x)}$

Comme $\mathcal{T}_{\vec{a}}$ est unitaire, $\mathcal{T}_{\vec{a}}^{\dagger} = \mathcal{T}_{\vec{a}}^{-1}$; en multipliant membre à membre à gauche par $\mathcal{T}_{\vec{a}\,\perp\,\mathrm{O}z}$, on obtient :

$$\boxed{\forall\ \vec{a} \perp \mathrm{O}z\ :\quad H\mathcal{T}_{\vec{a}\,\perp\,\mathrm{O}z} = \mathcal{T}_{\vec{a}\,\perp\,\mathrm{O}z}H \quad\Longleftrightarrow\quad [H,\, \mathcal{T}_{\vec{a}\,\perp\,\mathrm{O}z}] = 0} \tag{17.74}$$

D'où le résultat majeur : le Hamiltonien *commute* avec les opérateurs de symétrie interne du système, une affirmation qui est quasi-évidente après coup puisque par nature, le système est totalement insensible à toute opération de symétrie. De fait, si $\vec{a}$ est perpendiculaire à $\mathrm{O}z$ ($\vec{a} = a_x\vec{e}_x + a_y\vec{e}_y$), l'opérateur de translation donné en (17.68) devient :

$$\mathcal{T}_{\vec{a}\,\perp\,\mathrm{O}z} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}(a_xp_x + a_yp_y)}\ ; \tag{17.75}$$

comme H ne contient ni x ni y – les deux seules coordonnées affectées par p_x et p_y –, H et $\mathcal{T}_{\vec{a}\,\perp\,\mathrm{O}z}$ commutent.

Ainsi, le Hamiltonien commute avec tout opérateur associé à une opération de symétrie propre du système. Le Hamiltonien étant le générateur du mouvement, toute propriété de symétrie (la parité, la façon de se transformer par une rotation, etc.) est une "constante du mouvement"[44]. À titre d'exemple, montrons explicitement qu'un état pair ne peut se transformer en état impair par évolution libre[45]. Soit la valeur moyenne $\langle\Pi\rangle \stackrel{\text{déf}}{=} \langle\Psi|\Pi|\Psi\rangle$ calculée avec un certain état ; par l'équation de Schrödinger, on a :

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\Pi\rangle = -\frac{1}{\mathrm{i}\hbar}\langle\Psi|H\Pi|\Psi\rangle + \frac{1}{\mathrm{i}\hbar}\langle\Psi|\Pi H|\Psi\rangle \equiv \frac{1}{\mathrm{i}\hbar}\langle[\Pi,\, H]\rangle\ . \tag{17.76}$$

Si H est invariant par parité, le commutateur $[\Pi,\, H]$ est nul et la valeur moyenne $\langle\Pi\rangle$ est une constante dans le temps, quel que soit l'état $|\Psi\rangle$: l'opérateur Π – qui est linéaire – est une constante du mouvement. Pour un état de parité déterminée $|\Psi_\varepsilon\rangle$ ($\Pi|\Psi_\varepsilon\rangle = \varepsilon|\Psi_\varepsilon\rangle$, $\varepsilon = \pm 1$), $\langle\Pi\rangle = \varepsilon$; pour un état quelconque, $|\Psi\rangle \stackrel{\text{déf}}{=} \cos\theta|\Psi_{+1}\rangle + \mathrm{e}^{\mathrm{i}\alpha}\sin\theta|\Psi_{-1}\rangle$, la moyenne $\langle\Pi\rangle$, toujours constante dans le temps, vaut $\cos 2\theta$.

Puisqu'ils commutent, H et les opérateurs liés à la symétrie interne possèdent des vecteurs propres en commun, ce qui facilitera grandement la résolution effective de l'équation aux valeurs et vecteurs propres. Rappelons que l'existence de vecteurs propres communs ne vaut pas seulement pour des opérateurs hermitiques : si A et B hermitiques commutent, ils ont les mêmes vecteurs propres, en effet ; mais l'opérateur $\mathrm{e}^{\mathrm{i}B}$, qui n'est pas hermitique (il est unitaire), a les mêmes vecteurs propres que B.

---

[44] Rien n'interdit formellement de définir la moyenne quantique d'un opérateur qui n'est pas hermitique, en particulier celle de tout opérateur représentant une opération de symétrie ; la moyenne en question ne sera pas forcément réelle, puisque l'opérateur considéré est unitaire ou, exceptionnellement, antiunitaire, mais peu importe. De même, on peut écrire des équations de Heisenberg pour un opérateur non hermitique ; pour les opérateurs *linéaires* décrivant la symétrie, ces équations montrent qu'il s'agit de constantes du mouvement, précisément.

[45] On dit parfois de cette impossibilité qu'elle résulte d'une règle de "supersélection", pour l'opposer aux règles de sélection usuelles qui, le plus souvent, sont définies à l'intérieur d'une approximation donnée (pour l'interaction atome - champ, on parle ainsi de règles de sélection *dipolaires électriques* (E1), *quadrupolaires électriques* (E2), *dipolaires magnétiques* (M1), etc., voir chapitre 24, sous-section 24.5.1). Comme toujours, la symétrie fonctionne par tout ou rien et verrouille sans nuances.

♦ *Remarques*

1. Il convient de distinguer les deux types de conditions imposées par l'invariance vis-à-vis d'une opération de symétrie $\mathcal{S}$, qu'elle soit euclidienne ou de symétrie propre. S'agissant des probabilités, il faut bien :

$$|(|\Psi'\rangle,\, |\Phi'\rangle)|^2 = |(|\Psi\rangle,\, |\Phi\rangle)|^2 \ , \tag{17.77}$$

quels que soient les états $|\Psi\rangle$ et $|\Phi\rangle$; c'est bien le *module carré* qu'il faut considérer et, suivant le théorème de Wigner, ceci ouvre une double possibilité : $\mathcal{S}$ est soit unitaire, soit antiunitaire.

D'un autre côté, il est bien évident que toutes les valeurs moyennes doivent être elles aussi invariantes : si A est une observable quelconque, les valeurs moyennes de A doivent être les mêmes pour deux observateurs, liés à deux repères différents pour la symétrie euclidienne, ou acteur et aveugle en ce qui concerne la symétrie interne[46] ; ceci impose :

$$\langle\Psi'|A'|\Psi'\rangle = \langle\Psi|A|\Psi\rangle \qquad \forall\,|\Psi\rangle \ , \tag{17.78}$$

d'où il résulte[47] que les deux opérateurs $\mathcal{S}^\dagger A'\mathcal{S}$ et A sont égaux ; le transformé d'un opérateur linéaire A représentant une observable est donc :

$$\boxed{A' = \mathcal{S}A\mathcal{S}^\dagger} \tag{17.79}$$

La relation (17.79) étant établie, il en résulte, cette fois pour les éléments non-diagonaux (pas seulement pour les valeurs moyennes, qui sont des éléments diagonaux) :

$$(|\Psi\rangle,\, A|\Phi\rangle) = (|\Psi\rangle,\, \mathcal{S}^\dagger A'\mathcal{S}|\Phi\rangle) \qquad \forall|\Psi\rangle,\ |\Phi\rangle \ . \tag{17.80}$$

Maintenant de deux choses l'une :

- ou bien $\mathcal{S}$ est unitaire, alors le second membre de (17.80) est $(\mathcal{S}|\Psi\rangle,\, A'\mathcal{S}|\Phi\rangle)$, soit $(|\Psi'\rangle,\, A'|\Phi'\rangle)$ et on retrouve, même pour les éléments non-diagonaux, une relation du genre (17.71) :

$$(|\Psi'\rangle,\, A'|\Phi'\rangle) = (|\Psi\rangle,\, A|\Phi\rangle) \qquad \forall|\Psi\rangle,\ |\Phi\rangle \ . \tag{17.81}$$

- ou bien $\mathcal{S}$ est antiunitaire ; alors le second membre de (17.80), compte tenu de la définition (17.32) de l'adjoint d'un opérateur antiunitaire, est égal à $(\mathcal{S}|\Psi\rangle,\, A'\mathcal{S}|\Phi\rangle)^*$ et maintenant :

$$(|\Psi'\rangle,\, A'|\Phi'\rangle) = (|\Psi\rangle,\, A|\Phi\rangle)^* \qquad \forall|\Psi\rangle,\ |\Phi\rangle \ . \tag{17.82}$$

[46] L'observateur acteur est celui qui translate le système, l'autre est celui qui se ferme les yeux pendant la translation.

[47] voir [15], p. 549.

2. À titre d'exemple, revenons à la translation de vecteur $\vec{a}$, conduisant à la relation (17.63) entre les valeurs moyennes. Imaginons maintenant deux observateurs du même système : l'un ne bouge pas, l'autre suit le système dans sa translation ; ce dernier trouve la valeur moyenne de la coordonnée en manipulant l'opérateur transformé $\vec{r}\,'$ défini comme :

$$\vec{r}\,' = \mathcal{T}_{\vec{a}}\,\vec{r}\,\mathcal{T}_{\vec{a}}^{\dagger} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\vec{a}.\vec{p}}\,\vec{r}\,\big(\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\vec{a}.\vec{p}}\big)^{\dagger} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\vec{a}.\vec{p}}\,\vec{r}\,\mathrm{e}^{-\frac{1}{\mathrm{i}\hbar}\vec{a}.\vec{p}}\ , \tag{17.83}$$

conformément à (17.68) et (17.79). Le développement de chaque exponentielle fait apparaître une série de commutateurs multiples, série dont seuls les deux premiers termes sont non-nuls, puisque le premier commutateur est proportionnel à **1** :

$$\vec{r}\,' = \vec{r} + \frac{1}{\mathrm{i}\hbar}\,[\vec{a}.\vec{p}\,, \vec{r}\,] + \frac{1}{(\mathrm{i}\hbar)^2}\,[\vec{a}.\vec{p},\ [\vec{a}.\vec{p}, \vec{r}\,]] + \ldots = \vec{r} + \frac{1}{\mathrm{i}\hbar}\,[\vec{a}.\vec{p}\,, \vec{r}\,] = \vec{r} - \vec{a}\ . \tag{17.84}$$

L'observateur déplacé avec le système calcule la valeur moyenne par l'expression[48] $\langle\Psi'|\vec{r}\,'|\Psi'\rangle \stackrel{\text{déf}}{=} \langle\vec{r}\,'\rangle'$:

$$\langle\vec{r}\,'\rangle' = \int_{\mathbb{R}^3} \Psi^*(T_{\vec{a}}^{-1}\vec{r},\, t)(\vec{r}-\vec{a})\Psi(T_{\vec{a}}^{-1}\,\vec{r},\, t)\,\mathrm{d}^3r = \int_{\mathbb{R}^3} \Psi^*(\vec{\rho},\, t)\,\vec{\rho}\,\Psi(\vec{\rho},\, t)\,\mathrm{d}^3\rho\ , \tag{17.85}$$

l'intégrale de droite résultant du simple changement de variable $\vec{\rho} = \vec{r} - \vec{a}$; celle-ci n'est autre que la valeur moyenne du rayon-vecteur (peu importe le nom de la variable muette d'intégration) et, au total :

$$\boxed{\langle\vec{r}\,'\rangle' = \langle\vec{r}\,\rangle} \tag{17.86}$$

Cette identité exprime que les deux observateurs trouvent le même résultat chacun dans son repère, comme le requiert l'invariance euclidienne : que l'on soit à Paris ou à Papeete, on trouve la même longueur pour la liaison chimique H_2 !

On peut dire que dans ce type d'analyse, il y a en réalité trois observateurs ; deux d'entre eux sont liés chacun à un repère R ou R' (ici R' est translaté de $\vec{a}$ par rapport à R). Le troisième joue le rôle de *juge de paix* et compare les résultats des deux autres. Le premier observateur mesure $\langle\vec{r}\,\rangle$, le second mesure $\langle\vec{r}\,'\rangle'$ qui est identique à ce qu'a trouvé le premier. Le troisième observateur constate avec soulagement que l'invariance euclidienne imposée donne la bonne relation $\langle\vec{r}\,'\rangle = \langle\vec{r}\,\rangle + \vec{a}$ (voir (17.63)).

3. Les deux conditions (17.51) et (17.78) ne sont pas réductibles l'une à l'autre. En effet, même si les probabilités sont identiques (assurées par la première condition), les valeurs moyennes seraient différentes si les deux opérateurs A et A' n'avaient pas le même spectre de valeurs propres. Il est donc bien nécessaire d'imposer la deuxième condition ; c'est aussi suffisant car la relation (17.79) est une relation d'équivalence assurant que A et A' ont bel et bien le même spectre. Avec l'identité des probabilités et l'identité des spectres, l'invariance est assurée.

[48] Bien noter le $'$ supplémentaire à *l'intérieur* de $\langle\ldots\rangle$ qui représente la moyenne.

4. Examinons enfin ce que deviennent les relations de commutation dans une transformation $\mathcal{S}$, en raisonnant simplement avec x et p_x. Partant de[49] $[x,\, p_x] = \mathrm{i}\hbar\mathbf{1}$, multiplions à gauche par $\mathcal{S}$ et à droite par $\mathcal{S}^\dagger = \mathcal{S}^{-1}$:

$$\mathcal{S}\,[x,\, p_x]\,\mathcal{S}^\dagger = \mathcal{S}\,\mathrm{i}\hbar\,\mathcal{S}^\dagger \ . \tag{17.87}$$

Dans tous les cas, on a $\mathcal{S}\mathcal{S}^\dagger = \mathcal{S}^\dagger\mathcal{S} = \mathbf{1}$; ceci permet d'insérer le produit $\mathcal{S}\mathcal{S}^\dagger$ là où il faut, et on voit de suite que le premier membre est de fait toujours égal à $[x', p'_x]$ ($x' = \mathcal{S}x\mathcal{S}^\dagger$, $p'_x = \mathcal{S}p_x\mathcal{S}^\dagger$). Au contraire, le second membre dépend du caractère unitaire ou antiunitaire de $\mathcal{S}$. Si $\mathcal{S}$ est unitaire, on a : $\mathcal{S}(\mathrm{i}\hbar)\mathcal{S}^\dagger = \mathrm{i}\hbar\mathcal{S}\mathcal{S}^\dagger = \mathrm{i}\hbar\mathbf{1}$; si S est antiunitaire : $\mathcal{S}(\mathrm{i}\hbar)\mathcal{S}^\dagger = -\mathrm{i}\hbar\mathcal{S}\mathcal{S}^\dagger = -\mathrm{i}\hbar\mathbf{1}$.

Plus généralement, soit A et B deux observables ; étant hermitiques, leur commutateur peut être noté $\mathrm{i}C$ où C est aussi hermitique :

$$[A,\, B] = \mathrm{i}\,C \ ; \tag{17.88}$$

entre opérateurs transformés, on a la relation $\mathcal{S}\,[A,\, B]\,\mathcal{S}^\dagger = \mathcal{S}(\mathrm{i}\,C)\mathcal{S}^\dagger$, d'où :

$$\boxed{[A',\, B'] = \begin{cases} +\mathrm{i}\,\mathcal{S}C\mathcal{S}^\dagger = +\mathrm{i}C' & \text{si } \mathcal{S} \text{ est unitaire} \\ -\mathrm{i}\,\mathcal{S}C\mathcal{S}^\dagger = -\mathrm{i}C' & \text{si } \mathcal{S} \text{ est antiunitaire} \end{cases}} \tag{17.89}$$

Autrement dit, les relations de commutation fondamentales sont invariantes si la symétrie est unitaire et changent de signe si la symétrie est antiunitaire. Comme on le verra par la suite, le renversement du temps, représenté par la transformation notée K, est antiunitaire. Il en résultera les relations, nécessaires – et évidentes physiquement[50] :

$$\boxed{\vec{r}\,' = K\,\vec{r}\,K^\dagger = \vec{r} \ , \qquad \vec{p}\,' = K\,\vec{p}\,K^\dagger = -\,\vec{p}} \tag{17.90}$$

de sorte que la relation de commutation fondamentale $[u,\, p_v] = \mathrm{i}\hbar\delta_{uv}\,\mathbf{1}$ *change de signe* et devient :

$$\boxed{[u,'\, p'_v] = -\,\mathrm{i}\hbar\delta_{uv}\,\mathbf{1}} \tag{17.91}$$

♦

17.3.2 Rotations

Les arguments développés à propos des translations dans l'espace peuvent être repris pour les rotations. L'invariance par translation implique l'homogénéité présupposée de l'espace (supposé illimité), telle que l'affirme le principe euclidien (pas d'origine privilégiée) ; en ce qui concerne les rotations, c'est l'isotropie de l'espace qui est à l'œuvre (pas de direction

[49] $\mathbf{1}$ désigne l'opérateur identité, parfois omis quand aucune ambiguïté n'est possible.

[50] Changer t en $-t$ ne modifie pas les coordonnées mais inverse toutes les vitesses (et donc aussi les moments cinétiques).

privilégiée). Les rotations sont également des déplacements et, comme il existe des transformations infinitésimales, arbitrairement voisines de la transformation identité, les opérateurs associés aux rotations et agissant dans l'espace des états seront aussi unitaires. L'analyse qui suit jouera un rôle de tout premier plan dans l'étude des systèmes possédant une symétrie de rotation ; le champ central (voir chapitre 19) en fournira un exemple très important. Ces arguments d'invariance par rotation seront aussi utilisés à propos des effets Zeeman et Stark, permettant d'obtenir sans calcul des résultats qualitatifs importants sur la levée de dégénérescence consécutive à la brisure de la symétrie sphérique par le champ externe appliqué.

Considérons une rotation du système d'un angle θ autour d'un axe défini par le vecteur unitaire $\vec{u}$. Par définition, dans cette transformation $R_{\vec{u},\,\theta}$, $\vec{r}$ devient $\vec{r}\,'$:

$$\boxed{\vec{r}\,' \stackrel{\text{déf}}{=} R_{\vec{u},\,\theta}\,\vec{r}} \tag{17.92}$$

Soit $\Psi(\vec{r},\,t)$ une fonction d'onde du système ; lorsque ce dernier subit la rotation $R_{\vec{u},\,\theta}$, la fonction d'onde devient $\Psi'(\vec{r},\,t)$, qui doit pouvoir se déduire de $\Psi(\vec{r},\,t)$ par l'action d'un opérateur $\mathcal{R}_{\vec{u},\,\theta}$; à nouveau, l'invariance euclidienne permet d'écrire :

$$\boxed{\Psi'(\vec{r},\,t) \stackrel{\text{déf}}{=} \mathcal{R}_{\vec{u},\,\theta}\,\Psi(\vec{r},\,t) = \Psi(R_{\vec{u},\,\theta}^{-1}\,\vec{r},\,t)} \tag{17.93}$$

qui est l'analogue de (17.58). L'argument conduisant à cette relation est exactement le même que pour les translations : l'invariance des probabilités requiert l'invariance des modules au carré des produits scalaires, ce qui autorise toutefois une phase additionnelle non-triviale entre Ψ' et Ψ. Cependant, cette phase (variable dans l'espace) donnerait une nouvelle absurdité : le moment cinétique du système – pendant de l'impulsion quand il s'agit de rotation – serait changé quand on fait tourner le système. Par ailleurs, l'unitarité prévue de $\mathcal{R}_{\vec{u},\,\theta}$ se retrouve de la même façon que pour une translation (il suffit à nouveau de changer les variables muettes d'intégration). Enfin, en transposant les arguments ayant conduit à (17.86), on montre sans peine que la valeur moyenne de la coordonnée dans l'état tourné est le transformé de la valeur moyenne dans l'état avant rotation par la rotation $\mathcal{R}_{\vec{u},\,\theta}$.

L'opérateur $\mathcal{R}_{\vec{u},\,\theta}$ est facile à trouver, en suivant une procédure du même type que pour les translations, c'est-à-dire en considérant d'abord une rotation d'angle infinitésimal. La variation $\delta\vec{r} \stackrel{\text{déf}}{=} \mathcal{R}_{\vec{u},\delta\theta}\,\vec{r} - \vec{r}$ est perpendiculaire à la fois à $\vec{u}$ et à $\vec{r}$; elle est donc colinéaire à $\vec{u}\times\vec{r}$, son module est $r_\perp\delta\theta$, avec $r_\perp = r\,|\sin(\vec{r},\vec{u})|$, d'où $\delta\vec{r} = \vec{u}\times\vec{r}\,\delta\theta$. Pour une rotation infinitésimale d'angle $\delta\theta$, le vecteur transformé de $\vec{r}$ est donc :

$$\vec{r}\,' \stackrel{\text{déf}}{=} R_{\vec{u},\delta\theta}\,\vec{r} = \vec{r} + \delta\theta\,\vec{u}\times\vec{r}\ . \tag{17.94}$$

Selon (17.93), on a[51] $\Psi'(\vec{r},\,t) = \Psi(\vec{r} - \delta\theta\,\vec{u}\times\vec{r},\,t)$, d'où :

$$\Psi'(\vec{r},\,t) = \Psi(\vec{r}) - \delta\theta\,\vec{u}\times\vec{r}.\vec{\nabla}\Psi(\vec{r},\,t) + \ldots \equiv (\mathbf{1} - \delta\theta\,\vec{u}.\vec{r}\times\vec{\nabla})\,\Psi(\vec{r},\,t) + \ldots\ . \tag{17.95}$$

[51]L'inverse d'une rotation d'angle θ est manifestement la rotation d'angle $-\theta$ autour du même axe.

En remplaçant $\vec{\nabla}$ par $-\vec{p}/(\mathrm{i}\hbar)$, on fait apparaître le moment cinétique orbital $\vec{L} = \vec{r} \times \vec{p}$:

$$\Psi'(\vec{r},\, t) = \mathcal{R}_{\vec{u},\delta\theta}\, \Psi(\vec{r},\, t) = \left(\mathbf{1} + \frac{\delta\theta}{\mathrm{i}\hbar}\, \vec{u}.\vec{L}\right) \Psi(\vec{r},\, t) + \ldots\ , \tag{17.96}$$

d'où par identification :

$$\boxed{\mathcal{R}_{\vec{u},\delta\theta} = \mathbf{1} + \frac{\delta\theta}{\mathrm{i}\hbar}\, \vec{u}.\vec{L} + \mathcal{O}(\delta\theta^2)} \tag{17.97}$$

Visiblement, $\vec{p}$ et $\vec{L}$ jouent des rôles analogues pour les translations et les rotations respectivement : $\vec{L}$ est ainsi appelé le générateur des rotations, comme le montre l'égalité (17.97), analogue de (17.66) pour les translations.

Pour obtenir l'opérateur associé à une rotation d'angle fini θ autour d'un axe donné, il sufit de combiner un grand nombre de rotations élémentaires : on pose $\delta\theta = \frac{\theta}{N}$ et on fait le produit de N rotations d'angle $\delta\theta$ autour du même[52] axe de rotation. On obtient ainsi, après passage à la limite $N \to \infty$, $\delta\theta \to 0$, et nouvelle prise en compte de $\lim_{N\to\infty}[1 + x/N + \mathcal{O}(N^{-2})]^N = \mathrm{e}^x$:

$$\boxed{\mathcal{R}_{\vec{u},\,\theta} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\, \vec{u}.\vec{L}}} \tag{17.98}$$

Cet opérateur est visiblement unitaire puisque le moment cinétique est hermitique.

Tout comme pour les rotations, on peut munir l'ensemble des rotations d'une structure de groupe. L'expression (17.98) corrobore le fait connu que deux rotations autour de deux axes *différents* (Ox et Oy, par exemple) ne commutent pas[53], puisque le commutateur $[L_x,\, L_y]$ n'est pas nul (il est égal à $\mathrm{i}\hbar L_z$, voir Tome I, problème 7.4.6 et éqs.(12.8), (12.9)). En conséquence, le groupe complet des rotations n'est pas abélien (comparer avec (17.70)) :

$$\mathcal{R}_{\vec{u},\,\theta} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta \sum_{v=x,y,z} L_v u_v} \neq \prod_{v=x,y,z} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\, L_v u_v} \equiv \prod_{v=x,y,z} \mathcal{R}_{\mathrm{O}v,\,\theta}\ . \tag{17.99}$$

Bien évidemment, le sous-groupe des rotations autour d'un axe donné $\vec{u}$ est abélien, puisque $\mathcal{R}_{\vec{u},\,\theta_1}\mathcal{R}_{\vec{u},\,\theta_2} = \mathcal{R}_{\vec{u},\,\theta_1+\theta_2}$.

Quand on fait tourner le système, le Hamiltonien H se transforme en H' ; l'invariance euclidienne de la Mécanique quantique, cette fois par rapport aux rotations, requiert à nouveau l'invariance des valeurs moyennes – et aussi des éléments de matrice

[52] S'en convaincre sur un exemple simple : que devient le point de coordonnées (1, 0, 0) dans le repère $(\vec{i},\, \vec{j},\, \vec{k})$ quand on lui applique successivement les deux rotations $R_{\vec{i},\,\pi/2}$ et $R_{\vec{k},\,\pi/2}$ dans un ordre ou dans un autre ?

[53] Ainsi, au contraire des translations, les rotations ne commutent pas en général. Il est intéressant de noter que translations et rotations sont de surcroît *topologiquement* inéquivalentes, ne serait-ce parce que les unes (translations) n'ont pas de points fixes. Cette différence qualitative n'empêche pas, quand le terrain est bien balisé (en particulier, non-pertinence des effets de bord), d'assimiler souvent localement rotations et translations, par exemple pour un "grand" système où la transformation peut toujours être considérée comme infinitésimale (ce qui justifie l'usage des conditions cycliques de Born – von Kármán), et à condition de prendre les limites dans le bon ordre (celui qui est physiquement pertinent).

puisque $\mathcal{R}_{\vec{u},\theta}$ est unitaire, voir Remarque 1 p. 678 ; en particulier, pour H :

$$(|\Psi'\rangle,\, H'|\Phi'\rangle) = (|\Psi\rangle,\, H|\Phi\rangle) \quad\Longleftrightarrow\quad \langle\Psi|\mathcal{R}^{\dagger}_{\vec{u},\,\theta}H'\mathcal{R}_{\vec{u},\,\theta}|\Phi\rangle = \langle\Psi|H|\Phi\rangle \qquad (17.100)$$

quels que soient $|\Psi\rangle$ et $|\Phi\rangle$, d'où $\mathcal{R}^{\dagger}_{\vec{u},\,\theta}H'\mathcal{R}_{\vec{u},\,\theta} = H$; compte tenu de l'unitarité de $\mathcal{R}_{\vec{u},\,\theta}$, H' est finalement donné par[54] :

$$\boxed{H' = \mathcal{R}_{\vec{u},\,\theta}H\mathcal{R}^{\dagger}_{\vec{u},\theta} = \mathcal{R}_{\vec{u},\,\theta}H\mathcal{R}^{-1}_{\vec{u},\theta} \qquad \text{(invariance euclidienne)}} \qquad (17.101)$$

Maintenant, si le système possède la symétrie (interne) de rotation autour de l'axe défini par $\vec{u}$, son Hamiltonien est invariant :

$$H' = H \qquad \text{(invariance par symétrie propre) ;} \qquad (17.102)$$

utilisant l'unitarité de $\mathcal{R}_{\vec{u},\,\theta}$, il vient :

$$\boxed{H = \mathcal{R}_{\vec{u},\,\theta}H\mathcal{R}^{\dagger}_{\vec{u},\,\theta} \quad\Longleftrightarrow\quad [H,\, \mathcal{R}_{\vec{u},\,\theta}] = 0} \qquad (17.103)$$

Cette relation est vraie, quels que soient $\vec{u}$ et θ lorsque H ne dépend que du module carré du moment cinétique et du module du rayon vecteur, r (champ central, voir chapitre 19) ; H étant à symétrie sphérique, les trois composantes du moment cinétique orbital commutent chacune avec H et c'est bien le *vecteur* moment cinétique $\vec{L}$ qui est une constante du mouvement : tout comme pour les translations ressort le lien indissoluble entre symétrie et conservation dans le temps. À l'inverse, en présence d'un champ (dirigé le long de Oz), si le terme d'énergie cinétique continue à n'être fonction que du module carré du moment cinétique (voir l'expression (19.9)), le terme potentiel sera seulement invariant par rotation autour de Oz[55] : le champ appliqué brise la symétrie sphérique.

La relation (17.103) conduit naturellement à la notion d'observable *scalaire*, généralisant la notion usuelle de scalaire en géométrie ordinaire[56]. On dit qu'une observable A est un opérateur scalaire ssi elle est invariante dans toute rotation : $A' = A$; cette propriété se traduit par :

$$\boxed{A \text{ est un scalaire} \Longleftrightarrow A = \mathcal{R}_{\vec{u},\,\theta}A\mathcal{R}^{\dagger}_{\vec{u},\,\theta} \Longleftrightarrow [A,\, \mathcal{R}_{\vec{u},\,\theta}] = 0 \quad \forall\, \vec{u},\, \forall\, \theta} \qquad (17.104)$$

Selon (17.98), ceci veut dire qu'un opérateur scalaire commute avec chacune des trois composantes du moment cinétique orbital :

$$\boxed{[A,\, L_x] = 0 \qquad [A,\, L_y] = 0 \qquad [A,\, L_z] = 0} \qquad (17.105)$$

Par exemple, pour un système invariant par rotation, le Hamiltonien est un opérateur scalaire, ainsi que le carré $\vec{L}^2$ du moment cinétique.

[54] Soit une charge q couplée à un champ électrique $\vec{\mathcal{E}}$ dirigé le long de Oz ; l'interaction est $-q\mathcal{E}z$. Faisons la rotation de $+\pi/2$ autour de Oy : l'axe Oz se transforme en Ox. On vérifie bien que (17.100) et $\mathcal{R}_{\mathrm{O}y,\,\pi/2} = \mathrm{e}^{(1/\mathrm{i}\hbar)(\pi/2)L_y}$ conduisent à $H' = -q\mathcal{E}x$ comme il se doit.

[55] et seule la composante du moment cinétique le long de Oz sera une constante du mouvement.

[56] Le produit scalaire de deux vecteurs ordinaires est un invariant dans toute rotation des deux vecteurs.

Dans le même ordre d'idée et par généralisation, on introduit les opérateurs dits *vectoriels* ; un opérateur vectoriel $\vec{A}$ est défini par la propriété que les *valeurs moyennes* (quantiques) de ses composantes se comportent par rotation comme les composantes d'un vecteur de la géométrie ordinaire ; soit R une rotation dans $\mathbb{R}^3$ induisant la transformation (unitaire) $\mathcal{R}$ dans l'espace des états : $|\Psi'\rangle = \mathcal{R}\,|\Psi\rangle$. Comme précédemment, on note $\langle A_u \rangle$ ($u = x,\, y,\, z$) la valeur moyenne dans l'état $|\Psi\rangle$ avant rotation et $\langle A_u \rangle'$ la valeur moyenne dans l'état $|\Psi'\rangle$ après rotation. Compte tenu de la définition adoptée, on doit avoir :

$$\begin{bmatrix} \langle A_x \rangle' \\ \langle A_y \rangle' \\ \langle A_z \rangle' \end{bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{bmatrix} \begin{bmatrix} \langle A_x \rangle \\ \langle A_y \rangle \\ \langle A_z \rangle \end{bmatrix} \tag{17.106}$$

où les R_{uv} sont les éléments de la matrice de la rotation en question dans l'espace à trois dimensions[57]. L'expression de l'opérateur de rotation est déjà connue, donnée par (17.98), de sorte que tout est calculable dans l'équation (17.106). Le calcul explicite met en évidence un ensemble de relations caractéristiques entre les composantes d'un opérateur réputé vectoriel et les composantes du moment cinétique orbital $\vec{L}$ (énoncées sous forme de commutateurs).

On peut aussi utiliser l'argument qui suit, exploitant directement la définition *physique* des opérateurs vectoriels adoptée ici – traduite par (17.106) – afin notamment de montrer comment le moment cinétique "sort" spontanément dès qu'il s'agit de rotations. Ce faisant, d'une part on établit ces relations de commutation caractéristiques, d'autre part on retrouve l'expression de $\mathcal{R}_{\vec{u},\,\theta}$, déjà obtenue par d'autres moyens.

Considérons plus précisément une transformation infinitésimale, par exemple une rotation de $\delta\theta$ autour de Oz ; dans ce cas, (17.106) est explicitement :

$$\begin{bmatrix} \langle A_x \rangle' \\ \langle A_y \rangle' \\ \langle A_z \rangle' \end{bmatrix} = \begin{bmatrix} 1 & -\delta\theta & 0 \\ +\delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \langle A_x \rangle \\ \langle A_y \rangle \\ \langle A_z \rangle \end{bmatrix} + \mathcal{O}(\delta\theta^2) \ . \tag{17.109}$$

L'opérateur de rotation étant unitaire, on peut *a priori* l'écrire :

$$\mathcal{R}_{z,\,\delta\theta} = \mathbf{1} \ - \ \mathrm{i}\,\Omega_z\,\delta\theta + \mathcal{O}(\delta\theta^2) \ , \tag{17.110}$$

où Ω_z est un opérateur hermitique et où l'indice z rappelle qu'il s'agit d'une rotation autour de Oz. Avec ceci, les valeurs moyennes deviennent :

$$\langle A_x \rangle' \equiv \langle \Psi' | A_x | \Psi' \rangle = \langle \Psi | \big(\mathbf{1} + \mathrm{i}\Omega_z\delta\theta + \mathcal{O}(\delta\theta^2)\big) A_x \big(\mathbf{1} - \mathrm{i}\,\Omega_z\delta\theta + \mathcal{O}(\delta\theta^2)\big) | \Psi \rangle \ ; \tag{17.111}$$

[57] On sait écrire cette matrice indépendamment du présent contexte ; par exemple, dans la rotation autour de l'axe Oz d'un angle θ, les composantes d'un vecteur $\vec{V}$ se transforment en :

$$V_x' = V_x \cos\theta - V_y \sin\theta \ , \qquad V_y' = V_x \sin\theta + V_y \cos\theta \ , \qquad V_z' = V_z \ , \tag{17.107}$$

ce qui s'écrit sous forme matricielle :

$$\begin{bmatrix} V_x' \\ V_y' \\ V_z' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} \ . \tag{17.108}$$

Avec $|\delta\theta| \ll \pi$, on obtient par développement limité la matrice figurant dans (17.109). L'ensemble des rotations de $\mathbb{R}^3$ est noté O(3).

où l'hermiticité de Ω_z a été utilisée ; après développement on obtient :

$$\langle A_x \rangle' = \langle A_x \rangle + \mathrm{i}\delta\theta \langle \Psi | [\Omega_z, A_x] | \Psi \rangle + \mathcal{O}(\delta\theta^2) \ . \tag{17.112}$$

D'après (17.109), ceci doit être égal à $\langle A_x \rangle - \delta\theta \langle A_y \rangle + \mathcal{O}(\delta\theta^2)$, $\forall\, |\Psi\rangle$, d'où par identification des termes en $\delta\theta$:

$$\mathrm{i}\,[\Omega_z, A_x] = -A_y \quad \Longleftrightarrow \quad [\Omega_z, A_x] = \mathrm{i}A_y \ . \tag{17.113}$$

En procédant de même avec les composantes suivant Oy et Oz, on trouve :

$$[\Omega_z, A_y] = -\mathrm{i}A_x \ , \qquad [\Omega_z, A_z] = 0 \ . \tag{17.114}$$

Effectuant le même travail pour des rotations infinitésimales autour de Ox et Oy, on obtient cete fois :

$$[\Omega_x, A_y] = \mathrm{i}A_z \ , \qquad [\Omega_x, A_z] = -\mathrm{i}A_y \ , \qquad [\Omega_x, A_x] = 0 \ , \tag{17.115}$$

$$[\Omega_y, A_z] = \mathrm{i}A_x \ , \qquad [\Omega_y, A_x] = -\mathrm{i}A_z \ , \qquad [\Omega_y, A_y] = 0 \ . \tag{17.116}$$

Ces relations sont vraies pour n'importe quel opérateur vectoriel $\vec{A}$ ainsi défini ; il reste à trouver le triplet $(\Omega_x, \Omega_y, \Omega_z)$. Physiquement, le moment cinétique $\vec{L}$ doit être lui-même un opérateur vectoriel : on attend des valeurs moyennes de ses composantes qu'elles se transforment comme les composantes cartésiennes d'un vecteur. Il en résulte que dans toutes les relations ci-dessus, on peut remplacer $\vec{A}$ par $\vec{L}$. L'écriture explicite des neuf relations distinctes (17.113) - (17.116) permet alors de constater que le vecteur $\vec{\Omega}$ n'est autre que le moment cinétique $\vec{L}$ lui-même – au facteur $\hbar$ près, $\vec{\Omega} = \hbar^{-1}\vec{L}$. Cette identification permet de reformuler comme suit les relations (17.113) - (17.116), et d'écrire les relations *caractéristiques* satisfaites, par définition, par un opérateur vectoriel $\vec{V}$:

$$\boxed{\begin{array}{lll} [L_x, V_y] = \mathrm{i}\hbar V_z \ , & [L_y, V_y] = 0 \ , & [L_z, V_y] = -\mathrm{i}\hbar V_x \\ [L_y, V_z] = \mathrm{i}\hbar V_x \ , & [L_z, V_z] = 0 \ , & [L_x, V_z] = -\mathrm{i}\hbar V_y \\ [L_z, V_x] = \mathrm{i}\hbar V_y \ , & [L_x, V_x] = 0 \ , & [L_y, V_x] = -\mathrm{i}\hbar V_z \end{array}} \tag{17.117}$$

Il est facile de vérifier que le produit scalaire $\vec{V}_1.\vec{V}_2$ de deux opérateurs vectoriels est un opérateur scalaire, satisfaisant (17.105), et que leur produit vectoriel est un opérateur vectoriel (voir problème 17.8.1).

En conclusion, et en vue de la généralisation aux opérateurs tensoriels irréductibles, on peut affirmer qu'il existe une totale équivalence entre les deux définitions :

- un opérateur vectoriel est un ensemble de trois opérateurs dont les valeurs moyennes des composantes se transforment par rotation comme les trois composantes d'un vecteur de la géométrie ordinaire.
- un opérateur vectoriel est un ensemble de trois opérateurs satisfaisant avec les composantes du moment cinétique les neuf équations caractéristiques (17.117), ou

les relations équivalentes[58] :

$$\boxed{\begin{array}{lll} [L_+,\, V_x] = +\hbar\, V_z \ , & [L_+,\, V_y] = +\mathrm{i}\hbar\, V_z \ , & [L_+,\, V_z] = -\hbar\, V_+ \\ [L_-,\, V_x] = -\hbar\, V_z \ , & [L_-,\, V_y] = +\mathrm{i}\hbar\, V_z \ , & [L_-,\, V_z] = +\hbar\, V_- \\ [L_z,\, V_x] = +\mathrm{i}\hbar\, V_y \ , & [L_z,\, V_y] = -\mathrm{i}\hbar\, V_x \ , & [L_z,\, V_z] = 0 \end{array}} \qquad (17.118)$$

Ces relations peuvent d'ailleurs s'écrire de façon plus systématique en introduisant les composantes standard $V_q^{(1)}$ de l'opérateur vectoriel $\vec{V} \equiv V^{(1)}$, définies comme :

$$V_1^{(1)} \stackrel{\text{déf}}{=} -\frac{1}{\sqrt{2}}(V_x + \mathrm{i}\, V_y) \ , \quad V_0^{(1)} \stackrel{\text{déf}}{=} V_z \ , \quad V_{-1}^{(1)} \stackrel{\text{déf}}{=} +\frac{1}{\sqrt{2}}(V_x - \mathrm{i}\, V_y) \ . \quad (17.119)$$

Avec cette définition, et en notant généralement $\vec{J}$ le moment cinétique, on a :

$$\boxed{[J_\pm,\, V_q^{(1)}] = \hbar\,\sqrt{2 - q(q\pm 1)}\; V_{q\pm 1}^{(1)} \qquad [J_z,\, V_q^{(1)}] = \hbar q\; V_q^{(1)}} \qquad (17.120)$$

Ce sont ces dernières relations que l'on généralise en définissant les opérateurs tensoriels irréductibles (voir chapitre 18, section 18.6). Avec les définitions (17.119), le produit scalaire $\vec{V}.\vec{W}$ de deux opérateurs vectoriels s'écrit :

$$\vec{V}.\vec{W} = V_z W_z + \frac{1}{2}(V_+ W_- + V_- W_+) = V_0^{(1)} W_0^{(1)} - (V_1^{(1)} W_{-1}^{(1)} + V_{-1}^{(1)} W_1^{(1)}) \ . \quad (17.121)$$

♦ *Remarques*

1. On a ici délibérément adopté une approche intuitive pour la définition d'un opérateur vectoriel, en raisonnant avec les moyennes d'un certain triplet, et en leur imposant, pour des raisons physiques, de se transformer comme les composantes d'un vecteur au sens de la géométrie élémentaire.

   S'agissant du spin (voir chapitre 20), on argumentera exactement de la même façon en s'appuyant sur le fait physique que le moment magnétique $\vec{\mu}$ doit être un opérateur vectoriel au sens défini comme ci-dessus (les valeurs moyennes de ses composantes se transforment comme les composantes cartésiennes d'un vecteur ordinaire). Comme $\vec{\mu}$ et le spin $\vec{S}$ sont proportionnels, cette réquisition physique pour $\vec{\mu}$ fait *ipso facto* de $\vec{S}$ un opérateur vectoriel. Quand on fait tourner l'aimant qui permet de mesurer l'une ou l'autre des composantes de $\vec{\mu}$ ou de $\vec{S}$, on induit une transformation $\mathcal{R}_{\text{spin}}$ agissant dans l'espace des états de spin. Les calculs précédents peuvent être recommencés pour obtenir l'expression de $\mathcal{R}_{\text{spin}}$; en conformité avec (17.98), on trouve :

$$\boxed{\mathcal{R}_{\text{spin}\,\vec{u},\,\theta} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,\vec{u}.\vec{S}}} \qquad (17.122)$$

   D'autres arguments seront d'ailleurs développés dans le chapitre 20, qui permettent de se convaincre que $\vec{S}$ est effectivement un moment cinétique.

[58] $L_\pm = L_x \pm \mathrm{i}L_y$, $V_\pm = V_x \pm \mathrm{i}V_y$.

2. En désignant d'une façon générale par $\vec{J}$ le moment cinétique, l'opérateur de rotation est $\mathcal{R}_{\vec{u},\,\theta} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,\vec{u}.\vec{J}}$, d'où, pour une rotation d'un angle 2π :

$$\mathcal{R}_{\vec{u},\,2\pi} = \mathrm{e}^{\frac{2\pi}{\mathrm{i}\hbar}\,\vec{u}.\vec{J}} \tag{17.123}$$

   A priori, une rotation de 2π est une opération blanche, et l'on est donc tenté d'en déduire que les seules valeurs admissibles pour le produit scalaire $\vec{u}.\vec{J}$ sont des valeurs *entières*. Cette conclusion est erronée : on verra ultérieurement (chapitre 18) que le seul jeu des relations de commutation du moment cinétique exige que ce produit scalaire prenne des valeurs entières *ou* demi-entières. Le fait est que l'espace vectoriel dont la Mécanique quantique a besoin est construit sur des complexes, et que seuls les *rayons* $\mathrm{e}^{\mathrm{i}\alpha}|\Psi\rangle$ importent : une phase induite par une opération "blanche" n'est donc jamais à exclure d'emblée.

   Les bizarreries du cas demi-entier seront rediscutées dans le chapitre 20. En tout cas, aucun des arguments précédents n'implique cette distinction ; il s'ensuit que tous les résultats obtenus sont vrais pour toutes les valeurs du moment cinétique, qui se révéleront nécessairement entières s'il s'agit d'un moment orbital, demi-entières quand interviennent des spins demi-entiers en nombre impair. La possibilité de valeurs demi-entières conduit à un phénomène surprenant à première vue : pour un spin $J = 1/2$, la rotation de 2π autour de $\vec{u}$ n'est pas l'identité. En effet, dans ce cas, l'égalité (17.123) donne $\mathcal{R}_{\vec{u},\,2\pi} = -\mathbf{1}$; cette particularité sera analysée dans le chapitre 20.

3. Comme discuté par Delamotte [12], la notion de *vecteur* suppose bien plus qu'une structure *linéaire* (un espace vectoriel, par définition) puisque c'est en définitive un objet dont les composantes se transforment linéairement et d'une façon bien précise par rotation dans un espace de dimension donnée, quelle que soit la nature précise de ce dernier. Dans le cas de $\mathbb{R}^3$, le mode de transformation est déterminé par les matrices orthogonales 3×3 de déterminant égal à $+1$ constituant le groupe des rotations SO(3). C'est pourquoi, d'une façon plus abstraite, un *vecteur* de $\mathbb{R}^3$ peut en fait être défini comme engendrant une *représentation* linéaire de ce groupe, lequel traduit l'isotropie de cet espace euclidien. Cette notion garde donc toute sa pertinence qu'il s'agisse de l'espace "réel" ou de celui des impulsions, ou de l'espace de Hilbert d'un système. Ainsi, dans un espace de dimension D, un vecteur est (par définition) un ensemble de "D-plet" (D composantes) engendrent une représentation pour le groupe SO(D). ♦

17.4 Invariance de jauge

Les éléments de symétrie d'un système ne sont pas toujours de nature géométrique. De fait, la notion d'invariance par rapport à certaines transformations peut aussi résulter de la formalisation d'un problème donné en termes d'objets qui n'ont pas en eux-mêmes une signification physique directe, même si les éléments de la réalité physique peuvent en être déduits.

C'est le cas de l'invariance de jauge, qui sera discutée ci-dessous pour le champ électromagnétique. Pour une particule de charge q soumise à un champ $(\vec{\mathcal{E}}, \vec{\mathcal{B}})$, on sait poser d'emblée les équations du mouvement à l'aide des champs, en écrivant l'équation fondamentale de la Dynamique où apparaît la force de Lorentz $\vec{F}_{\rm L} \stackrel{\text{déf}}{=} q(\vec{\mathcal{E}} + \vec{v} \times \vec{\mathcal{B}})$. D'un autre côté, il se révèle commode pour les calculs d'exprimer les champs eux-mêmes à l'aide des potentiels. Pour un même champ, on peut définir une infinité de potentiels se déduisant les uns des autres par ce que l'on appelle une transformation de jauge, et qui tous donnent les mêmes équations du mouvement en Mécanique classique : les potentiels dépendent de la jauge choisie mais les champs eux-mêmes – et donc les forces – y sont insensibles. L'invariance de jauge la plus triviale est la possibilité d'ajouter une constante additive quelconque à une énergie potentielle : $V(\vec{r})$ et $V(\vec{r}) + \mathrm{C}^{\rm ste}$ ont le même gradient, donc définissent la même force $\vec{F} = -\vec{\nabla} V$.

Les équations de Maxwell – qui donnent les champs, donc les forces agissant sur une particule chargée – sont invariantes quand on passe d'un certain potentiel $(U, \vec{A})$ à un autre, $(U', \vec{A}')$, moyennant une transformation appropriée. En effet, comme $\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$ et que $\vec{\nabla} \times \vec{\nabla} = \vec{0}$, on peut, sans changer le champ magnétique, ajouter à $\vec{A}$ le gradient de n'importe quelle fonction de l'espace et du temps, $\phi(\vec{r}, t)$:

$$\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A} = \vec{\nabla} \times [\vec{A} + \vec{\nabla}\phi(\vec{r}, t)] \equiv \vec{\nabla} \times \vec{A}' \ . \tag{17.124}$$

Par ailleurs, le champ électrique est :

$$\vec{\mathcal{E}} = -\vec{\nabla} U - \frac{\partial \vec{A}}{\partial t} \quad ; \tag{17.125}$$

si on veut qu'il ne doit pas non plus altéré quand on remplace $\vec{A}$ par $\vec{A}'$, il faut simultanément changer U en $U' = U - \frac{\partial \phi}{\partial t}$. Au total, pour toute fonction ϕ, le changement des potentiels :

$$\boxed{U(\vec{r}, t) \rightarrow U'(\vec{r}, t) \stackrel{\text{déf}}{=} U(\vec{r}, t) - \frac{\partial \phi}{\partial t} \ , \quad \vec{A}(\vec{r}, t) \rightarrow \vec{A}'(\vec{r}, t) \stackrel{\text{déf}}{=} \vec{A}(\vec{r}, t) + \vec{\nabla}\phi(\vec{r}, t)} \tag{17.126}$$

ne modifie par construction ni le champ électrique, ni le champ magnétique et porte le nom de *transformation de jauge*.

On sait par ailleurs que le Lagrangien L pour une particule de charge q située dans un champ associé à $(U, \vec{A})$ est (voir Tome I, éq. (7.18)) :

$$L = \frac{1}{2} m\vec{v}^{\,2} - qU(\vec{r}, t) + q\vec{A}(\vec{r}, t).\vec{v} \ ; \tag{17.127}$$

pour un autre choix de jauge, le Lagrangien sera donc L', différent de L ; si on désigne généralement par des primes ($'$) les grandeurs dynamiques pertinentes quand on travaille dans la deuxième jauge :

$$L' = \frac{1}{2} m\vec{v}^{\,\prime 2} - qU'(\vec{r}^{\,\prime}, t) + q\vec{A}'(\vec{r}^{\,\prime}, t).\vec{v}^{\,\prime} \ . \tag{17.128}$$

D'un autre côté, la vitesse et la position de la particule sont indépendantes de la jauge en Mécanique classique, puisqu'elles se déduisent directement de l'équation fondamentale $m\dot{\vec{v}} = \vec{F}_\mathrm{L}$, où les potentiels n'apparaissent pas : on dit que vitesse et position sont des quantités invariantes de jauge ; l'invariance de jauge de la position et de la vitesse en Mécanique classique s'exprime suivant :

$$\boxed{\vec{r}\,' = \vec{r} \qquad m\vec{v}\,' = m\vec{v} \qquad \text{(classique)}} \tag{17.129}$$

Le moment de la quantité de mouvement, $\vec{r} \times m\vec{v}$ est donc lui aussi invariant ; au contraire, le moment conjugué $\vec{p} \overset{\text{déf}}{=} \vec{\nabla}_{\vec{v}} L$, lui, n'est pas invariant de jauge ; selon (17.127) et (17.128), on a respectivement :

$$\vec{p} \overset{\text{déf}}{=} \frac{\partial L}{\partial \vec{v}} = m\vec{v} + q\vec{A}(\vec{r},\, t) \ , \qquad \vec{p}\,' \overset{\text{déf}}{=} \frac{\partial L'}{\partial \vec{v}\,'} = m\vec{v}\,' + q\vec{A}'(\vec{r},\, t) \tag{17.130}$$

mais comme la vitesse est invariante ($\vec{v}\,' = \vec{v}$) :

$$\vec{p}\,' = m\vec{v} + q\vec{A}'(\vec{r},\, t) \ , \tag{17.131}$$

et au total :

$$\boxed{\vec{p}\,' = \vec{p} + q[\vec{A}'(\vec{r},\, t) - \vec{A}(\vec{r},\, t)] \neq \vec{p} \qquad \text{(classique)}} \tag{17.132}$$

L'écart entre les moments conjugués dans les deux jauges n'implique que le potentiel-vecteur. Noter que la présence d'un tel champ exige de distinguer entre les moments cinétiques $\vec{r} \times (m\vec{v})$ et $\vec{r} \times \vec{p}$; dans le cadre classique, le premier est invariant de jauge, le second ne l'est pas.

♦ *Remarques*

1. Comme mentionné plus haut, l'adjonction d'une constante additive au potentiel scalaire est le cas le plus élémentaire d'invariance de jauge. De fait, avec $\vec{A}(\vec{r},\, t) \equiv 0$ et $U(\vec{r})$ (indépendant du temps), le champ magnétique est nul et le champ électrique est constant dans le temps : c'est la situation standard de l'Électrostatique. Faisons maintenant une transformation de jauge selon (17.126) avec une fonction ϕ indépendante de $\vec{r}$ et linéaire en temps ($\phi = -Kt$) ; dans ces conditions, $\vec{A}' \equiv \vec{0}$ (le champ magnétique reste nul) et U' ne diffère de U que par la constante additive K : $U'(\vec{r}) = U(\vec{r}) + K$.

   Il en résulte que les deux Lagrangiens L et L' ne diffère que par la constante K ; l'intégration en temps donnant l'action S fournira donc le seul terme additionnel $K(t_2 - t_1)$, qui disparaîtra lors de la variation et conduira aux mêmes équations de Lagrange.

   Noter d'ailleurs que l'on peut aussi prendre une fonction $\phi(t)$ *quelconque*. En effet, quel que soit ϕ ne dépendant que du temps, les deux Lagrangiens ne diffèrent que par une dérivée totale en temps, et donnent les mêmes équations de Lagrange (cet argument a été déjà utilisé : voir la dérivation du Lagrangien pour une particule libre, p. 662, et notamment l'équation (17.15)). En définitive, pour n'importe quel problème, les deux énergies potentielles $V(\vec{r})$ et $V'(\vec{r},\, t) \overset{\text{déf}}{=} V(\vec{r}) + \phi(t)$ donnent exactement les mêmes équations du mouvement.

2. Si le Lagrangien n'est pas invariant de jauge (puisque U et $\vec{A}$ y apparaissent explicitement), les équations de Lagrange le sont : L a été précisément construit pour reproduire l'équation de Newton avec la force de Lorentz au second membre.

   En revanche, les équations de Hamilton ne sont pas invariantes de jauge (tout comme le Hamiltonien), notamment parce que $\vec{p} \neq m\vec{v}$. Dans les notations précédentes, on a :

$$H = \frac{1}{2m}[\vec{p} - q\vec{A}(\vec{r}, t)]^2 + qU(\vec{r}, t) \ ; \qquad (17.133)$$

   la première équation de Hamilton est $\dot{\vec{r}} = \vec{\nabla}_{\vec{p}} H = \frac{1}{m}(\vec{p} - q\vec{A})$, la seconde est $\dot{\vec{p}} = \vec{F}_{\rm L} + q\frac{\partial \vec{A}}{\partial t}$, en conformité avec $m\dot{\vec{v}} = \vec{F}_{\rm L}$ et $\vec{p} = m\vec{v} + q\vec{A}$. ♦

Examinons maintenant comment se traduit l'invariance de jauge en Mécanique quantique où, comme on va le voir, les conclusions sont en quelque sorte inversées. Les grandeurs dynamiques ($\vec{r}$, $\vec{v}$, $\vec{p}$, etc.) sont maintenant représentées par des opérateurs et l'état dynamique du système l'est par le vecteur d'état $|\Psi\rangle$. La transformation de jauge est une opération de symétrie, il existe donc un opérateur $\mathcal{G}$ reliant les deux vecteurs d'états représentant le système dans une jauge ou une autre :

$$U(\vec{r}, t) \rightarrow U'(\vec{r}, t) \ , \vec{A}(\vec{r}, t) \rightarrow \vec{A}'(\vec{r}, t) \quad \Longleftrightarrow \quad |\Psi\rangle \rightarrow |\Psi'\rangle \stackrel{\text{déf}}{=} \mathcal{G}|\Psi\rangle \ . \qquad (17.134)$$

$\mathcal{G}$ est sûrement unitaire (pas antiunitaire), car il existe des transformations de jauge infiniment proches de la transformation identité (il suffit de prendre une fonction $\phi(\vec{r}, t)$ arbitrairement petite – ce qui revient, formellement, à faire tendre la charge électrique vers zéro, voir (17.148)). Continuant à mettre un prime ($'$) pour les grandeurs associées à la jauge (U', $\vec{A}'$), l'invariance requise par la symétrie de jauge s'exprime par l'égalité des valeurs moyennes de la position et de la vitesse (ou, de façon équivalente, de la quantité de mouvement). *A priori*, on doit donc écrire :

$$\langle\vec{r}'\rangle' = \langle\vec{r}\rangle \Longleftrightarrow \int_{\mathbb{R}^3} \Psi'^*(\vec{r}, t)\vec{r}'\Psi'(\vec{r}, t)\, \mathrm{d}^3r = \int_{\mathbb{R}^3} \Psi^*(\vec{r}, t)\vec{r}\,\Psi(\vec{r}, t)\, \mathrm{d}^3r \ , \qquad (17.135)$$

$$\langle m\vec{v}'\rangle' = \langle m\vec{v}\rangle \Longleftrightarrow \int_{\mathbb{R}^3} \Psi'^*(\vec{r}, t)m\vec{v}'\Psi'(\vec{r}, t)\, \mathrm{d}^3r = \int_{\mathbb{R}^3} \Psi^*(\vec{r}, t)m\vec{v}\,\Psi(\vec{r}, t)\, \mathrm{d}^3r \ . \qquad (17.136)$$

C'est bien ainsi que doit s'exprimer la symétrie de jauge en Mécanique quantique, et non par l'invariance des opérateurs $\vec{r}$ et $\vec{v}$ qui n'ont pas *a priori* un sens physique direct – seules les *valeurs moyennes* quantiques en ont un. D'ailleurs et de toute façon, il est clair que les conditions (17.135) et (17.136) doivent être vraies dans la limite classique – laquelle s'exprime précisément en termes de valeurs *moyennes*, à comparer le cas échéant aux grandeurs mécaniques classiques.

À l'inverse, les relations canoniques de quantification ont été énoncées en toute généralité, indépendamment de toute référence à une jauge particulière. Il en résulte que, en représentation-q, on doit *dans toutes les jauges* utiliser la correspondance universelle de Schrödinger :

$$\boxed{\vec{r} \longrightarrow \times\vec{r} \qquad \vec{p} \longrightarrow -\mathrm{i}\hbar\vec{\nabla}} \qquad (17.137)$$

ce qui assure le maintien de la relation fondamentale $[q,\, p] = \mathrm{i}\hbar\mathbf{1}$, et ceci quelle que soit la jauge. Il en résulte que les opérateurs $\vec{r}$ et $\vec{p}$ sont invariants de jauge par construction de la théorie ; par conséquent les opérateurs primés coïncident avec les opérateurs non primés (comparer avec (17.129)) :

$$\boxed{\vec{r}\,' = \vec{r} \qquad \vec{p}\,' = \vec{p} \qquad \text{(quantique)}} \tag{17.138}$$

En revanche, pour l'opérateur associé à la quantité de mouvement $m\vec{v}$, on a (compte tenu de (17.138)) :

$$m\vec{v} = \vec{p} - q\vec{A}(\vec{r},\, t) \ , \qquad m\vec{v}\,' = \vec{p}\,' - q\vec{A}'(\vec{r},\, t) = \vec{p} - q\vec{A}'(\vec{r},\, t) \ . \tag{17.139}$$

soit (comparer avec (17.132)) :

$$\boxed{m\vec{v}\,' = m\vec{v} + q[\vec{A}(\vec{r},\, t) - \vec{A}'(\vec{r},\, t)] \neq m\vec{v} \qquad \text{(quantique)}} \tag{17.140}$$

En conséquence, l'opérateur associé à la quantité de mouvement $m\vec{v}$ n'est pas invariant de jauge, ce qui n'est pas problématique puisque les prévisions ne s'expriment pas à l'aide des seuls opérateurs[59] mais impliquent simultanément le vecteur d'état.

En vertu de (17.138) et (17.139), les équations d'invariance (17.135) et (17.136) s'expriment alors comme :

$$\langle\Psi'|\vec{r}\,|\Psi'\rangle = \langle\Psi|\vec{r}\,|\Psi\rangle \ , \qquad \langle\Psi'|\vec{p} - q\vec{A}'|\Psi'\rangle = \langle\Psi|\vec{p} - q\vec{A}\,|\Psi\rangle \ . \tag{17.141}$$

Avec la définition (17.134) de la transformation $\mathcal{G}$, les relations (17.141), vraies $\forall\ |\Psi\rangle$, peuvent être élevées au rang d'égalités entre opérateurs :

$$\boxed{\mathcal{G}^\dagger\,\vec{r}\,\mathcal{G} = \vec{r} \qquad \mathcal{G}^\dagger\,(\vec{p} - q\vec{A}')\,\mathcal{G} = \vec{p} - q\vec{A}} \tag{17.142}$$

On en déduit que la transformation $\mathcal{G}$ ne dépend que de $\vec{r}$ (elle ne peut dépendre de $\vec{p}$, qui changerait forcément $\vec{r}$). $\mathcal{G}$ commute donc avec les potentiels-vecteurs, ce qui permet de transformer (17.142) comme suit (et tenant compte de (17.126)) :

$$\mathcal{G}^\dagger\,\vec{p}\,\mathcal{G} - q\vec{A}' = \vec{p} - q\vec{A} \quad \Longleftrightarrow \quad \mathcal{G}^\dagger\,\vec{p}\,\mathcal{G} = \vec{p} + q\vec{\nabla}\phi \ . \tag{17.143}$$

$\mathcal{G}$, fonction de $\vec{r}$, est un opérateur unitaire ; on peut donc toujours l'écrire :

$$\mathcal{G} = \mathrm{e}^{\,\mathrm{i}F(\vec{r},\,t)} \ , \tag{17.144}$$

où F est une fonction (un scalaire) à valeurs réelles. Dans ces conditions, la deuxième écriture de (17.143) devient :

$$\mathrm{e}^{-\mathrm{i}F(\vec{r},\,t)}\,\vec{p}\,\mathrm{e}^{\,\mathrm{i}F(\vec{r},\,t)} = \vec{p} + q\vec{\nabla}\phi \ . \tag{17.145}$$

[59]Le spectre d'un opérateur donne les valeurs possibles des résultats d'une mesure : il doit donc être invariant de jauge. C'est bien le cas puisque la relation $\Omega' = \mathcal{G}\Omega\mathcal{G}^\dagger$ est une relation d'équivalence assurant que les deux opérateurs Ω' et Ω ont les mêmes valeurs propres.

Le premier membre est :

$$e^{-iF(\vec{r},\,t)}\,\vec{p}\,e^{iF(\vec{r},\,t)} = \vec{p} - i[F(\vec{r},\,t),\,\vec{p}] + \frac{i^2}{2}\,[F(\vec{r},\,t),\,[F(\vec{r},\,t),\,\vec{p}]] + \ldots \;; \qquad (17.146)$$

comme $[F(\vec{r},\,t),\,\vec{p}] = i\hbar\vec{\nabla}F(\vec{r},\,t)\mathbf{1}$, le commutateur d'ordre 2 est nul ainsi que tous les suivants. L'équation (17.143) devenant $\vec{p} + \hbar\,\vec{\nabla}F = \vec{p} + q\vec{\nabla}\phi$, son intégration membre à membre donne :

$$F(\vec{r},\,t) = \frac{q}{\hbar}\,\phi(\vec{r},\,t) + F_0(t) \;. \qquad (17.147)$$

La fonction $F_0(t)$, apparaissant en tant que constante de l'intégration par rapport à $\vec{r}$, donne un facteur de phase global indépendant de la coordonnée à l'état transformé, et peut donc être omise ; finalement, la fonction F à retenir est[60] :

$$F(\vec{r},\,t) = \frac{q}{\hbar}\,\phi(\vec{r},\,t) \;, \qquad (17.148)$$

et, selon (17.134), à l'état transformé $|\Psi'\rangle$ correspond la fonction d'onde :

$$\boxed{\Psi'(\vec{r},\,t) \stackrel{\text{déf}}{=} \mathcal{G}\Psi(\vec{r},\,t) = e^{\frac{i}{\hbar}q\phi(\vec{r},\,t)}\,\Psi(\vec{r},\,t)} \qquad (17.149)$$

Bien noter que le facteur de phase, parce qu'il dépend de l'espace, n'est pas trivial et ne peut de ce fait être ignoré. S'il y a non pas une mais N particules de charges $q_1, q_2, \ldots, q_N$, la transformation de jauge entre fonctions d'onde s'écrit :

$$\Psi'(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N,\,t) = e^{\frac{i}{\hbar}\sum_n q_n\phi(\vec{r}_n,\,t)}\,\Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N,\,t) \;. \qquad (17.150)$$

Montrons maintenant que l'équation de Schrödinger est invariante de jauge, ce qui est bien la moindre des choses, puisqu'elle a été induite sans aucune restriction vis-à-vis d'un éventuel champ électromagnétique. Pour cela, il faut que les deux équations suivantes soient vérifiées :

$$i\hbar\,\frac{\partial}{\partial t}\,|\Psi\rangle = H|\Psi\rangle \quad (\text{S}) \;, \qquad i\hbar\,\frac{\partial}{\partial t}\,|\Psi'\rangle = H'|\Psi'\rangle \quad (\text{S'}) \;, \qquad (17.151)$$

avec :

$$|\Psi'\rangle = e^{iF}|\Psi\rangle \;, \quad H = \frac{1}{2m}\,(\vec{p} - q\vec{A})^2 + qU \;, \quad H' = \frac{1}{2m}\,(\vec{p} - q\vec{A}')^2 + qU' \;. \qquad (17.152)$$

Après report de $|\Psi'\rangle = e^{iF}|\Psi\rangle$, le premier membre de (S') se développe comme suit :

$$i\hbar\,\frac{\partial}{\partial t}\,e^{iF}|\Psi(t)\rangle = i\hbar\Big(i\frac{\partial F}{\partial t}\,e^{iF}|\Psi(t)\rangle + e^{iF}\,\frac{\partial}{\partial t}\,|\Psi(t)\rangle\Big) \;, \qquad (17.153)$$

d'où une autre forme pour l'équation (S') :

$$i\hbar\,\frac{\partial}{\partial t}\,|\Psi\rangle = e^{-iF}\left(H' + \hbar\,\frac{\partial F}{\partial t}\right)e^{iF}\,|\Psi\rangle \;. \qquad (17.154)$$

[60] Remarquer que la fonction F contient la charge électrique en facteur : on peut donc formellement, faisant tendre celle-ci vers zéro, aller continûment vers la transformation de jauge identité.

Il faut que ceci soit identique à l'équation (S), contrainte qui, compte tenu de (17.148), se traduit par l'égalité :

$$\mathrm{e}^{-\mathrm{i}F}\Big(H' + q\,\frac{\partial\phi}{\partial t}\Big)\mathrm{e}^{\mathrm{i}F} = H \quad ; \tag{17.155}$$

or, d'après (17.146) et (17.126) :

$$\mathrm{e}^{-\mathrm{i}F}(\vec{p} - q\vec{A}')\mathrm{e}^{\mathrm{i}F} = \mathrm{e}^{-\mathrm{i}F}\vec{p}\,\mathrm{e}^{\mathrm{i}F} - q\vec{A}' = \vec{p} + q\vec{\nabla}\phi - q\vec{A}' = \vec{p} - q\vec{A} \quad ; \tag{17.156}$$

par ailleurs :

$$\mathrm{e}^{-\mathrm{i}F}\Big(qU' + q\frac{\partial\phi}{\partial t}\Big)\mathrm{e}^{\mathrm{i}F} = qU' + q\frac{\partial\phi}{\partial t} = qU \quad ; \tag{17.157}$$

donc l'équation (17.155) est vérifiée, assurant que (S') est équivalente à (S), c'est-à-dire que les deux équations (S) et (S') sont simultanément vraies : l'équation de Schrödinger, ne faisant référence à aucune jauge particulière, est bien invariante.

Pour terminer, montrons que la densité $\rho = \Psi^*\Psi$ et le courant associé, $\vec{j}$, sont invariants de jauge, ce qui n'est pas surprenant puisque l'équation de Schrödinger elle-même contient l'équation de conservation. En ce qui concerne ρ, c'est évident puisque la phase $\mathrm{e}^{\mathrm{i}F}$ disparaît dans le module au carré :

$$\boxed{\rho' = \rho} \tag{17.158}$$

Pour établir l'invariance du courant, il faut et suffit de prendre garde au fait qu'il s'exprime à l'aide de la *vitesse*[61] $\vec{v} = \frac{1}{m}(\vec{p} - q\vec{A})$; en présence d'un tel champ, le vecteur $\vec{j}$ a donc pour expression :

$$\vec{j} = \frac{1}{2}(\Psi^*\vec{v}\,\Psi + \Psi\vec{v}\,\Psi^*) \equiv \Re(\Psi^*\vec{v}\,\Psi) = \frac{1}{m}\Re[\Psi^*(\vec{p} - q\vec{A})\Psi] = \frac{1}{m}\Re[\Psi^*(-\mathrm{i}\hbar\vec{\nabla} - q\vec{A})\Psi] \ . \tag{17.159}$$

Dans l'autre jauge, on a :

$$\vec{j}' = \frac{1}{m}\Re[\Psi'^*(-\mathrm{i}\hbar\vec{\nabla} - q\vec{A}')\Psi'] = \frac{1}{m}\Re\left[\mathrm{e}^{-\mathrm{i}F}\Psi^*(-\mathrm{i}\hbar\vec{\nabla} - q\vec{A} - q\vec{\nabla}\phi)\mathrm{e}^{+\mathrm{i}F}\Psi\right] \ . \tag{17.160}$$

Le terme entre crochets à droite vaut $\Psi^*[-\mathrm{i}\hbar\mathrm{e}^{-\mathrm{i}F}(\mathrm{i}\vec{\nabla}F + \vec{\nabla}\Psi)\,\mathrm{e}^{+\mathrm{i}F} - q\vec{A} - q\vec{\nabla}\phi]\,\Psi$, d'où, compte tenu de (17.147) :

$$\vec{j}' = \frac{1}{m}\Re\left\{\Psi^*\left[-\mathrm{i}\hbar\mathrm{e}^{-\mathrm{i}F}(\mathrm{i}\hbar^{-1}q\vec{\nabla}\phi + \vec{\nabla}\Psi)\mathrm{e}^{+\mathrm{i}F} - q\vec{A} - q\vec{\nabla}\phi\right]\Psi\right\} \ , \tag{17.161}$$

soit :

$$\boxed{\vec{j}' = \frac{1}{m}\Re\left[\Psi^*(\vec{p} - q\vec{A})\Psi\right] \equiv \vec{j}} \tag{17.162}$$

Ainsi, la densité ρ et le courant associé $\vec{j}$, reliés par $\partial_t\rho + \vec{\nabla}.\vec{j} = 0$ sont chacun séparément invariants de jauge, en conformité avec le théorème de Noether.

[61] Tout comme l'énergie cinétique, qui vaut $(\vec{p} - q\vec{A})^2/(2m)$.

17.5 Symétries discrètes

Toutes les symétries rencontrées jusqu'à présent sont associées à des transformations continues, au sens où celles-ci dépendent continûment d'un paramètre (angle de rotation, amplitude de la translation, charge électrique pour la transformation de jauge, etc.). En conséquence, il existe des transformations infinitésimales, aussi proches que l'on veut de la transformation identité ; ceci assure que les opérateurs associés ne sauraient être antiunitaires : par le théorème de Wigner, ils sont donc forcément unitaires.

Il existe aussi des opérations de symétrie *discrètes* ; par exemple, pour un polygone régulier à N côtés, les rotations de $2\pi/N$ autour d'un axe perpendiculaire au plan et passant par le centre du polygone laissent celui-ci globalement invariant[62] ; le groupe correspondant est visiblement un sous-groupe (fini, lui) des rotations quelconques autour du même axe et est constitué d'opérateurs unitaires, quoique relatifs à des transformations discrètes. Pour un réseau régulier infini de points, il existe de même des translations finies superposant le réseau à lui-même, sous-groupe (infini dénombrable) du groupe (continu) des translations quelconques.

Il en va également ainsi de la réflexion par rapport à un plan (symétrie-miroir) qui, avec l'identité, forme un groupe à deux éléments. Il n'existe pas alors de transformation infinitésimale, ni de générateurs. Clairement, la réflexion change la chiralité et il est manifestement impossible de passer d'une forme chirale à une autre par continuité, sauf à inclure d'autres types de transformations, non considérées ici[63]. Un autre exemple est l'inversion d'espace, traditionnellement appelée parité en Mécanique quantique.

Ces transformations discrètes sont encore représentées par des opérateurs unitaires. Par exemple, pour la parité (inversion d'espace), représentée par l'opérateur noté Π agissant dans l'espace des états, la fonction d'onde se transforme suivant :

$$\boxed{\Psi'(\vec{r},\, t) \stackrel{\text{déf}}{=} \Pi\Psi(\vec{r},\, t) = \Psi(-\vec{r},\, t)} \tag{17.163}$$

Les arguments d'invariance galiléenne déjà invoqués, $\langle\vec{r}\,'\rangle' = \langle\vec{r}\rangle$ et $\langle\vec{p}\,'\rangle' = \langle\vec{p}\rangle$, montrent que les opérateurs position et moment conjugué changent de signe :

$$\boxed{\vec{r}\,' \equiv \Pi\,\vec{r}\,\Pi^\dagger = -\vec{r}\ , \qquad \vec{p}\,' \equiv \Pi\,\vec{p}\,\Pi^\dagger = -\vec{p}} \tag{17.164}$$

de sorte que les relations de commutation sont inchangées et Π est unitaire – ce que l'on peut tout autant vérifier en constatant l'invariance du produit scalaire (il suffit de faire un changement de variable muette dans les intégrales). Dans cette opération, le moment cinétique orbital $\vec{r}\times\vec{p}$ ne change pas de signe non plus. Si le Hamiltonien est lui-même insensible à Π, on peut trouver des vecteurs propres commun à H et Π ; il s'ensuit que, en l'absence de dégénérescence, toute fonction propre $\psi(\vec{r})$ de H est *nécessairement* paire *ou* impaire[64].

[62] Ce ne sont pas les seules symétries d'une telle figure.

[63] La pyramide NH_3 peut évidemment passer d'une forme à l'autre en s'aplatissant, mais ceci fait appel à des déformations (continues).

[64] Cet argument a été utilisé à plusieurs reprises, voir par exemple l'oscillateur harmonique à une dimension (Tome I, chapitre 16), et vaut pour tous les potentiels $V(x)$ pairs, $V(-x) = V(x)$.

Les mêmes observations valent pour la symétrie-miroir[65], représentée par $\mathcal{M}$; si $\vec{n}$ est le vecteur unitaire normal au plan de réflexion, on a :

$$\vec{r}\,' \equiv \mathcal{M}\,\vec{r}\,\mathcal{M}^\dagger = \vec{r} - 2(\vec{n}.\vec{r}\,)\,\vec{n} \ , \qquad \vec{p}\,' \equiv \mathcal{M}\,\vec{p}\,\mathcal{M}^\dagger = \vec{p} - 2(\vec{n}.\vec{p}\,)\,\vec{n} \tag{17.165}$$

et il est à nouveau facile de voir que les relations de commutation sont toujours inchangées (voir problème 17.8.2), d'où l'on déduit que $\mathcal{M}$ est unitaire. Dans la symétrie-miroir par rapport au plan $x\mathrm{O}z$, le report des transformés (17.165) dans $\vec{L}\,' = \vec{r}\,' \times \vec{p}\,'$ donne :

$$L'_x = -L_x \ , \qquad L'_y = L_y \ , \qquad L'_z = -L_z \ . \tag{17.166}$$

Physiquement, il est clair que si le point tourne autour de Oz, la symétrie-miroir par rapport au plan $x\mathrm{O}z$ inverse le sens de rotation (et de même si la rotation a lieu autour de Ox) ; en revanche, rien ne change si la rotation a lieu autour de l'axe Oy, qui est normal au plan-miroir. Au total, quoique discrètes, ces dernières symétries sont encore représentées par des opérateurs unitaires.

Il en va tout autrement pour une transformation importante, le renversement du temps, une opération qui n'est pas propre à la Mécanique quantique, et se présente déjà dans le cadre classique. Dans ce dernier cas, soit une particule de masse m soumise à la force $\vec{F}(\vec{r}\,) = -\vec{\nabla}V(\vec{r}\,)$, et définissons précisément l'opération de renversement du temps comme suit :

$$t_{\mathrm{renv}} = -t \ , \qquad \vec{r}_{\mathrm{renv}} = \vec{r} \ , \qquad \vec{v}_{\mathrm{renv}} = -\vec{v} \ . \tag{17.167}$$

Supposons que, dans l'équation de Newton (17.2), le second membre ne dépend que de la position $\vec{r}$; en conséquence, elle garde la même forme si on l'exprime en termes de toutes les grandeurs renversées, puisque $\mathrm{d}/\mathrm{d}t_{\mathrm{renv}} = -\mathrm{d}/\mathrm{d}t$. Schématiquement, cette invariance résulte du fait que l'équation est du second ordre par rapport au temps : comme le temps figure "au carré", changer t en $-t$ ne fait donc rien. Il en résulte que, pour des conditions initiales données $\vec{r}_0$ et $\vec{v}_0$, si la fonction $\vec{r}(t)$ est *la* solution des équations du mouvement classique, la fonction définie par :

$$\boxed{\vec{r}_{\mathrm{renv}}(t) \stackrel{\mathrm{déf}}{=} \vec{r}(-t)} \tag{17.168}$$

est aussi solution des mêmes équations avec des conditions initiales renversées (position inchangée, vitesse renversée).

L'invariance par renversement du temps peut s'exprimer également d'autres façons ; $\vec{r}$ est une certaine fonction, g, paramétrée par les conditions initiales (voir (17.8)). Si on inverse les temps et les vitesses – en laissant évidemment les coordonnées inchangées – on retrouve la même coordonnée $\vec{r}$; d'où, reprenant la notation introduite en (17.8) :

$$g(-t, -t_0\,;\, \vec{r}_0, -\vec{v}_0) = g(t, t_0\,;\, \vec{r}_0, \vec{v}_0) \ , \tag{17.169}$$

[65]La symétrie-miroir est aussi le produit d'une inversion et d'une rotation de π autour d'un axe perpendiculaire au miroir ; ce simple fait assure l'unitarité de $\mathcal{M}$.

égalité que l'on vérifie facilement dans des cas simples[66]. Enfin, l'invariance s'exprime aussi en disant que si la particule part du point A à $t = 0$ avec une vitesse $\vec{v}_{\rm A}$ et arrive en B à l'instant t avec la vitesse $\vec{v}_{\rm B}$, alors, partant de B à $t = 0$ avec la vitesse $-\vec{v}_{\rm B}$, elle est en A à l'instant t avec la vitesse $-\vec{v}_{\rm A}$. Les paramètres dynamiques (position, vitesse,...) sont associés les uns avec les autres de façon à respecter l'invariance $t \rightarrow -t$; par exemple, l'expression de la coordonnée ne peut contenir que des fonctions paires du produit vitesse×temps.

L'équation de Newton n'est pas toujours invariante par renversement du temps ; si la particule (chargée) est soumise à un champ magnétique, la force au second membre de (17.2) est la force de Lorentz, qui dépend de la vitesse : c'est pourquoi un champ magnétique brise l'invariance $t \rightarrow -t$. Il en va de même si on introduit une force de frottement (fluide par exemple), dépendant (forcément) de la vitesse.

La symétrie éventuelle par renversement du temps dépend généralement des propriétés de parité du Hamiltonien par rapport au moment conjugué $\vec{p}$. En l'absence de champ magnétique et de force de frottement, et pour un champ de force indépendant du temps, le Hamiltonien est $H(\vec{r}, \vec{p}) = \frac{\vec{p}^2}{2m} + V(\vec{r})$; comme $\vec{p}$ n'apparaît qu'au carré, le Hamiltonien H est *pair* en $\vec{p}$: $H(\vec{r}, \vec{p}) = H(\vec{r}, -\vec{p})$. Dans ces conditions, les équations de Hamilton $\dot{\vec{r}} = \vec{\nabla}_{\vec{p}} H$ et $\dot{\vec{p}} = -\vec{\nabla}_{\vec{r}} H$ sont invariantes quand on effectue la transformation définie en (17.167). En effet, la première équation devient :

$$\dot{\vec{r}}_{\rm renv} = -\dot{\vec{r}} = -\vec{\nabla}_{\vec{p}} H(\vec{r}, \vec{p}) = +\vec{\nabla}_{\vec{p}_{\rm renv}} H(\vec{r}_{\rm renv}, -\vec{p}_{\rm renv}) = +\vec{\nabla}_{\vec{p}_{\rm renv}} H(\vec{r}_{\rm renv}, \vec{p}_{\rm renv}) \ ; \tag{17.170}$$

quant à la deuxième équation, elle s'écrit (puisque $\vec{p}_{\rm renv} = -\vec{p}$) :

$$\dot{\vec{p}}_{\rm renv} = (-)^2 \dot{\vec{p}} = -\vec{\nabla}_{\vec{r}} H(\vec{r}, \vec{p}) = -\vec{\nabla}_{\vec{r}_{\rm renv}} H(\vec{r}_{\rm renv}, -\vec{p}_{\rm renv}) = -\vec{\nabla}_{\vec{r}_{\rm renv}} H(\vec{r}_{\rm renv}, \vec{p}_{\rm renv}) \ ; \tag{17.171}$$

à chaque fois, la dernière égalité vient du fait que H est une fonction paire de $\vec{p}$.

• Considérons maintenant la situation quantique où, comme on va le voir, le renversement du temps revient (en l'absence du spin[67]) à laisser invariantes les coordonnées et à inverser les vitesses. Changeons t en $-t$ dans l'équation de Schrödinger :

$$\mathrm{i}\hbar \frac{\partial}{\partial(-t)} \Psi(\vec{r}, -t) = [-\frac{\hbar^2}{2m} \Delta + V(\vec{r})] \, \Psi(\vec{r}, -t) \ . \tag{17.172}$$

Pour récupérer le signe + au premier membre, il suffit d'en prendre le complexe conjugué ; par ailleurs, V est une fonction à valeurs réelles[68], et le Laplacien Δ est aussi réel. Au

[66] Par exemple, pour un oscillateur harmonique linéaire, on a (notations évidentes)

$$x(t) = x_0 \cos\omega(t - t_0) + \frac{v_0}{\omega} \sin\omega(t - t_0) \ ,$$

expression qui est inaltérée si on renverse les temps et la vitesse initiale.

[67] Dans le chapitre 20, on reviendra sur le renversement du temps pour expliquer les modifications rendues nécessaires par l'existence du spin (voir section 20.4).

[68] Dans le cas contraire, H ne serait pas hermitique, l'évolution par $U(t) \stackrel{\rm déf}{=} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar} Ht}$ ne serait pas unitaire, on perdrait de la probabilité, etc.

total, en prenant les complexes conjugués des deux membres, on trouve :

$$\mathrm{i}\hbar\,\frac{\partial}{\partial t}\,\Psi^*(\vec{r}, -t) = [-\frac{\hbar^2}{2m}\,\Delta + V(\vec{r})]\,\Psi^*(\vec{r}, -t)\ . \qquad (17.173)$$

Ainsi, avec le Hamiltonien apparaissant au second membre de (17.173), si $\Psi(\vec{r}, t)$ est solution, la fonction $\Psi^*(\vec{r}, -t)$ est aussi solution. Ceci conduit naturellement à définir la fonction $\Psi_{\mathrm{renv}}(\vec{r}, t)$, transformée de Ψ par renversement du temps :

$$\boxed{\Psi_{\mathrm{renv}}(\vec{r}, t) \stackrel{\text{déf}}{=} \Psi^*(\vec{r}, -t)} \qquad (17.174)$$

Dans le cas d'un état stationnaire $\Psi_{n\,\mathrm{st}}(\vec{r}, t) \stackrel{\text{déf}}{=} \psi_n(\vec{r})\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\,E_n t}$, on a[69] :

$$\boxed{\Psi_{n\,\mathrm{st,\,renv}}(\vec{r}, t) \equiv \psi_n^*(\vec{r})\,\mathrm{e}^{\frac{1}{-\mathrm{i}\hbar}\,E_n(-t)} = \psi_n^*(\vec{r})\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\,E_n t} = \Psi^*_{n\,\mathrm{st}}(\vec{r}, -t)} \qquad (17.175)$$

et, pour une fonction indépendante du temps :

$$\psi_{n\,\mathrm{renv}}(\vec{r}) = \psi_n^*(\vec{r})\ . \qquad (17.176)$$

Dans la suite, on désigne par K l'opérateur effectuant le renversement du temps dans l'espace des états ; dans ces notations, par définition :

$$\boxed{|\Psi_{\mathrm{renv}}(t)\rangle \stackrel{\text{déf}}{=} K|\Psi(t)\rangle} \qquad (17.177)$$

Il est bien clair que K ne peut être un opérateur unitaire, à cause de la conjugaison complexe qui apparaît dans la définition (17.174) (voir aussi [4], p. 407). De fait, raisonnant par exemple sur la base propre de H, on voit que l'opérateur K change chaque état propre en son complexe conjugué (selon (17.176)) et conjugue donc tout produit scalaire : pour cette raison, K est un opérateur *antiunitaire* (voir (17.38)), qui toutefois ne change pas la norme d'un vecteur.

Une fois définie l'opération K, il convient d'examiner la transformation des observables. Comme précédemment, il existe une symétrie universelle traduisant le fait que, dans une théorie *mécanique*, le sens du temps est arbitraire[70], et purement conventionnel. Toutes les valeurs moyennes obtenues quand un choix a été fait doivent coïncider avec celles trouvées avec le choix contraire[71]. Il convient donc d'écrire, pour toute observable A transformée par K :

$$\langle A'\rangle' = \langle A\rangle \quad \Longleftrightarrow \quad (|\Psi'\rangle,\, A'|\Psi'\rangle) = (|\Psi\rangle,\, A|\Psi\rangle) \quad \forall|\Psi\rangle\ , \qquad (17.178)$$

soit, en utilisant (17.177) et en tenant compte de la définition (17.32) de l'opérateur adjoint d'un opérateur antiunitaire :

$$(K|\Psi\rangle, A'K|\Psi\rangle) = (|\Psi\rangle, A|\Psi\rangle) \quad \Longleftrightarrow \quad (|\Psi\rangle, K^\dagger A'K|\Psi\rangle)^* = (|\Psi\rangle, A|\Psi\rangle)\ . \qquad (17.179)$$

[69] Les énergies sont réelles !

[70] Le Second Principe est d'essence *statistique* et ne s'impose qu'après avoir renoncé, parce que c'est impossible et parce que ce serait absurde méthodologiquement parlant au vu de l'expérience, à décrire *tous* les degrés de liberté d'un système macroscopique.

[71] Tout comme le sens positif sur l'axe Ox, choisi par convention pour faire des calculs, n'a évidemment aucun sens physique par lui-même dans l'espace supposé homogène, isotrope, pair, etc.

A étant une observable, sa valeur moyenne est réelle ; la dernière égalité donne donc après conjugaison complexe des deux membres :

$$(|\Psi\rangle,\, K^\dagger A' K|\Psi\rangle) = (|\Psi\rangle,\, A|\Psi\rangle) \quad \forall\, |\Psi\rangle \quad \Longleftrightarrow \quad K^\dagger A' K = A \ . \tag{17.180}$$

Puisque K est antiunitaire, $K^\dagger = K^{-1}$ et l'on retrouve, même pour le cas antiunitaire, la relation habituelle pour la transformation des observables :

$$\boxed{A' = K\, A\, K^\dagger} \tag{17.181}$$

assurant que les prévisions physiques ne sont pas tributaires d'une convention de calcul.

Soit maintenant à trouver plus particulièrement les transformés des grandeurs fondamentales $\vec{r}$ et $\vec{p}$. Pour trouver $\vec{r}\,'$ et $\vec{p}\,'$, plaçons-nous en représentation coordonnée. K ne modifie pas la position[72], $\vec{r}$ est donc invariant :

$$\boxed{\vec{r}\,' = K\vec{r}K^\dagger = \vec{r} \quad \Longleftrightarrow \quad [\vec{r},\, K] = 0} \tag{17.182}$$

En ce qui concerne l'impulsion, on a :

$$\vec{p}\,' = K\,(-\mathrm{i}\hbar\vec{\nabla})\,K^\dagger = +\mathrm{i}\hbar\, K\,\vec{\nabla}\,K^\dagger \ ; \tag{17.183}$$

noter le changement de signe venant de la commutation de i avec K antiunitaire. Maintenant, quelle que soit la fonction $\Psi(\vec{r},\, t)$:

$$K\,\vec{\nabla}\Psi(\vec{r},\, t) = K\,\frac{\partial\Psi(\vec{r},\, t)}{\partial\vec{r}} = \Big(\frac{\partial\Psi(\vec{r},\, -t)}{\partial\vec{r}}\Big)^* = \vec{\nabla}\Psi^*(\vec{r},\, -t) = \vec{\nabla}\, K\Psi(\vec{r},\, t) \ , \tag{17.184}$$

autrement dit $[K,\, \vec{\nabla}] = 0$ (d'ailleurs, K ne modifie pas la coordonnée) ; en faisant commuter K et $\vec{\nabla}$ dans (17.183) et en utilisant $KK^\dagger = \mathbf{1}$, il vient finalement $\vec{p}\,' = +\mathrm{i}\hbar\,\vec{\nabla}$, soit :

$$\boxed{\vec{p}\,' = K\vec{p}K^\dagger = -\vec{p} \quad \Longleftrightarrow \quad \{\vec{p},\, K\} = 0} \tag{17.185}$$

qui donne le résultat attendu ; remarquer que $\vec{p}$ *anticommute* avec K. En conséquence, le moment cinétique orbital est lui aussi changé en son opposé :

$$\vec{L}\,' \equiv K\vec{L}K^\dagger = K(\vec{r}\times\vec{p})K^\dagger = (K\vec{r}K^\dagger)\times(K\vec{p}K^\dagger) = \vec{r}\times(-\vec{p}) \equiv -\vec{L} \ ; \tag{17.186}$$

Les égalités (17.182) et (17.185) entraînent que les relations de commutation pour les variables transformées *changent de signe*, signature de l'antiunitarité de l'opérateur K :

$$\boxed{[r'_u, p'_v] = -\mathrm{i}\hbar\delta_{uv}\mathbf{1}} \tag{17.187}$$

Cela étant, par construction de K et conformément à (17.178), les moyennes sont invariantes :

$$\boxed{\langle\vec{r}\,'\rangle' = \langle\vec{r}\rangle \ , \qquad \langle\vec{p}\,'\rangle' = \langle\vec{p}\rangle \ , \qquad \langle\vec{L}\,'\rangle' = \langle\vec{L}\rangle \ , \ldots} \tag{17.188}$$

[72] Dans la définition (17.174), c'est le même argument $\vec{r}$ qui figure dans les deux membres. En représentation-p on a aussi $K\vec{r}K^\dagger = K\mathrm{i}\hbar\frac{\partial}{\partial\vec{p}}K^\dagger = -\mathrm{i}\hbar\frac{\partial}{\partial(-\vec{p})}KK^\dagger = \mathrm{i}\hbar\frac{\partial}{\partial\vec{p}} = \vec{r}$.

H est une observable parmi d'autres et se transforme en $H' = KHK^\dagger$; compte tenu de (17.185), le Hamiltonien est invariant s'il ne contient que des puissances paires de $\vec{p}$: en représentation-q, H est alors un opérateur différentiel *réel*. Ce n'est pas le cas en présence d'un champ magnétique, où des termes du genre $\vec{A}.\vec{p}$ apparaissent – ce sont d'ailleurs eux qui produisent le couplage sous la forme $-\vec{\mu}.\vec{\mathcal{B}}$, proportionnel à $-\vec{L}.\vec{\mathcal{B}}$ par l'intermédiaire du facteur gyromagnétique ($\vec{\mu} \propto \vec{L}$, voir Tome I, problème 7.4.1) ; or, précisément, le moment cinétique est inversé par K, d'où la non-invariance de H. Physiquement, ceci est bien évident : tout comme dans le cas classique, un champ magnétique implique des boucles de courant sous une forme ou sous une autre, où interviennent des vitesses, lesquelles changent de signe par K. En revanche, un champ électrique préserve l'invariance $t \to -t$, puisque le couplage au champ n'implique que $\vec{r}$ et que H reste pair en $\vec{p}$. Le lien entre symétrie et dégénérescence sera précisé dans la section 17.6, mais il est possible d'ores et déjà d'entrevoir qu'un champ électrique lèvera *a priori* moins de dégénérescence qu'un champ magnétique. Au total, et *mutatis mutandis*, un champ magnétique a le même effet en quantique et en classique sur la symétrie par renversement du temps.

Quand H est invariant, $H' = H$ et la relation de commutation $[H,\, K] = 0$ se trouve *ipso facto* vérifiée ; toutefois, en raison du caractère antiunitaire de K, cela vaut la peine de s'en convaincre explicitement. Soit un état stationnaire d'énergie E_n, $\Psi_{n,\,\mathrm{st}}(\vec{r},\, t) = \mathrm{e}^{E_n t/(\mathrm{i}\hbar)}\,\psi_n(\vec{r})$, satisfaisant :

$$H\Psi_{n,\,\mathrm{st}}(\vec{r},\, t) = E_n\,\psi_n(\vec{r})\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}E_n t} \equiv E_n\,\Psi_{n,\,\mathrm{st}}(\vec{r},\, t)\ . \tag{17.189}$$

On a d'abord :

$$HK\Psi_{n,\,\mathrm{st}} = H\,\psi_n^*\mathrm{e}^{-\frac{1}{\mathrm{i}\hbar}E_n(-t)} = H\,\psi_n^*\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}E_n t}\ ; \tag{17.190}$$

comme H est réel – c'est pour ceci qu'il est invariant par K –, si on a $H\psi_n = E_n\psi_n$, on a également[73], en prenant les complexes conjugués des deux membres, $H\psi_n^* = E_n\psi_n^*$. Il vient donc :

$$HK\Psi = E_n\,\psi_n^*\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}E_n t}\ . \tag{17.191}$$

Ensuite, pour le produit KH :

$$KH\Psi = KE_n\,\psi_n\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}E_n t} = E_n\,\psi_n^*\mathrm{e}^{\frac{1}{-\mathrm{i}\hbar}E_n(-t)} = E_n\,\psi_n^*\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}E_n t}\ . \tag{17.192}$$

(17.191) et (17.192) montrent que HK et KH donnent le même résultat. La commutation de H avec K vaut tout autant pour une combinaison linéaire d'états stationnaires.

Bien évidemment, appliquer deux fois consécutivement le renversement du temps doit n'avoir aucune conséquence observable. Ceci entraîne que l'on doit avoir :

$$K^2|\Psi\rangle = \mathrm{e}^{\mathrm{i}\alpha}|\Psi\rangle\ , \tag{17.193}$$

[73]Noter au passage que pour un Hamiltonien invariant par renversement du temps, ψ_n et ψ_n^* correspondent à la même énergie propre. Dès lors, si le niveau E_n est non-dégénéré, on peut toujours supposer que la fonction ψ_n est réelle (une phase triviale peut toujours être ignorée) ; si le niveau est dégénéré, on peut toujours, si besoin est, former deux combinaisons linéaires réelles $(\psi_n + \psi_n^*)/\sqrt{2}$ et $\mathrm{i}(\psi_n^* - \psi_n)/\sqrt{2}$.

où α une phase qui est la même pour tous les états. On peut alors écrire successivement, pour deux états quelconques $|\Phi\rangle$ et $|\Psi\rangle$:

$$(K|\Phi\rangle,\, |\Psi\rangle) = (K|\Phi\rangle,\, K^\dagger K|\Psi\rangle) \qquad \text{puisque } K^\dagger K = \mathbf{1} \ , \tag{17.194}$$

et :

$$(K|\Phi\rangle,\, K^\dagger K|\Psi\rangle) = (K^2|\Phi\rangle,\, K|\Psi\rangle)^* \ , \tag{17.195}$$

par définition de l'adjoint d'un opérateur antilinéaire, d'où :

$$(K|\Phi\rangle,\, |\Psi\rangle) = (K^2|\Phi\rangle,\, K|\Psi\rangle)^* = (K|\Psi\rangle,\, K^2|\Phi\rangle) = (K|\Psi\rangle, \mathrm{e}^{\mathrm{i}\alpha}|\Phi\rangle) = \mathrm{e}^{\mathrm{i}\alpha}(K|\Psi\rangle, |\Phi\rangle) \ , \tag{17.196}$$

où (17.193) a été prise en compte. Les deux vecteurs $|\Phi\rangle$ et $|\Psi\rangle$ sont quelconques : on peut récrire (17.196) en les échangeant, soit $(K|\Psi\rangle,\, |\Phi\rangle) = \mathrm{e}^{\mathrm{i}\alpha}(K|\Phi\rangle, |\Psi\rangle)$; rapprochant cette égalité de (17.196), on obtient :

$$(K|\Phi\rangle,\, |\Psi\rangle) = \mathrm{e}^{\mathrm{i}\alpha}(K|\Psi\rangle, |\Phi\rangle) = \mathrm{e}^{2\mathrm{i}\alpha}\,(K|\Phi\rangle, |\Psi\rangle) \ , \tag{17.197}$$

d'où $\mathrm{e}^{2\mathrm{i}\alpha} = 1$, soit $\alpha = 0$ ou $\pi\,(2\pi)$. En d'autres termes, pour un système donné (et en fait selon la *nature* de ce système), on a :

$$\boxed{K^2|\Psi\rangle = |\Psi\rangle \quad ou \quad K^2|\Psi\rangle = -\,|\Psi\rangle} \tag{17.198}$$

Ces relations ont des conséquences importantes ; notamment, prenant $|\Phi\rangle = |\Psi\rangle$ dans (17.196), on trouve l'égalité :

$$(K|\Psi\rangle,\, |\Psi\rangle) = (K^2|\Psi\rangle,\, K|\Psi\rangle)^* = (K|\Psi\rangle,\, K^2|\Psi\rangle) = \mathrm{e}^{\mathrm{i}\alpha}(K|\Psi\rangle,\, |\Psi\rangle) \ , \tag{17.199}$$

de sorte que dans le cas $\alpha = \pi$:

$$\boxed{(K|\Psi\rangle,\, |\Psi\rangle) = -\,(K|\Psi\rangle,\, |\Psi\rangle) \quad \Longleftrightarrow \quad (K|\Psi\rangle,\, |\Psi\rangle) = 0} \tag{17.200}$$

Prenons pour $|\Psi\rangle$ l'un des états propres du système, $|\psi_n\rangle$; l'égalité (17.200) dit que $|\psi_n\rangle$ et $K|\psi_n\rangle$ sont orthogonaux ; comme, pour un Hamiltonien invariant par renversement du temps, un état propre et son renversé correspondent à la *même* valeur propre, il en résulte que lorsque $\alpha = \pi$, les énergies sont dégénérées au moins deux fois, chaque doublet étant formé des deux états (orthogonaux) qui se correspondent l'un l'autre par renversement du temps. Cette association par paires d'états dégénérés, globalement invariantes dans le changement $t \to -t$, porte le nom de *dégénérescence de Kramers.* Elle existe pour tous les systèmes à spin (total) demi-entier ; compte tenu du théorème d'addition des moments cinétiques (chapitre 18, notamment (18.204)), la dégénérescence de Kramers se produit pour les sytèmes à nombre *impair* de fermions (voir Remarque 1, p. 755).

♦ *Remarque*

Quand on revient à la définition première de K, (17.174) et (17.177), on ne voit pas bien comment K^2 peut être autre chose que l'identité, puisque conjuguer et/ou

changer t en $-t$ sont des opérations binaires. En fait, cette définition du renversement du temps doit être généralisée afin de tenir compte du spin[74]. On verra de fait que la phase α dépend du nombre $N_{1/2}$ de particules de spin demi-entier et vaut précisément $N_{1/2}\,\pi$ de sorte que :

$$\boxed{K^2 = \mathrm{e}^{\mathrm{i}N_{1/2}\pi}\,\mathbf{1} = (-1)^{N_{1/2}}\,\mathbf{1}} \tag{17.201}$$

En définitive, c'est le spin (et plus particulièrement le spin demi-entier) qui apporte quelque étrangeté et donne une phase pour une opération que l'intuition élémentaire aurait tendance à considérer comme *blanche* – et qui l'est effectivement pour des spins entiers (en particulier quand le spin est nul). On verra que c'est ce même spin qui, quand des nombres demi-entiers sont en jeu, introduit un changement de signe lors d'une rotation de 2π, une opération qui, dans ce cas non plus, n'est pas neutre (voir section 20.3, et notamment (20.111)). ♦

17.6 Symétrie et dégénérescence

L'existence d'une symétrie implique généralement l'existence d'états différents par certains de leurs attributs (c'est justement ce qui les distingue les uns des autres), mais devant posséder les mêmes propriétés observables puisque, par exemple, un système à symétrie sphérique se superpose à lui-même dans toute rotation : le système transformé et le système avant transformation sont indiscernables l'un de l'autre. Ainsi, l'existence d'une symétrie se traduit aussi nécessairement par l'apparition d'une *dégénérescence.*

Soit un système possédant une symétrie vis-à-vis de transformations R_i, géométriques ou non, représentées par des opérateurs unitaires ou antiunitaires notés $\mathcal{R}_i$. D'après ci-dessus, on a par hypothèse :

$$H' = \mathcal{R}_i H \mathcal{R}_i^\dagger = H \quad \Longleftrightarrow \quad [H, \mathcal{R}_i] = 0 \ . \tag{17.202}$$

Soit maintenant $|\psi_n\rangle$ un état propre de H associé à l'énergie E_n. Quand on applique R_i à un système situé dans l'état $|\psi_n\rangle$, ce vecteur se transforme en un vecteur $|\psi'_n\rangle$, *a priori* différent de $|\psi_n\rangle$ et tel que $|\psi'_n\rangle = \mathcal{R}_i\,|\psi_n\rangle$, par définition de l'opérateur $\mathcal{R}_i$ représentant la transformation R_i. Montrons maintenant que $|\psi'_n\rangle$ est encore propre de H, associé à la même énergie E_n, en accord avec le fait que les deux états $|\psi'_n\rangle$ et $|\psi_n\rangle$ étant image et objet l'un de l'autre dans une opération de symétrie, leurs énergies doivent être les mêmes. En effet, comme $[H,\, \mathcal{R}_i] = 0$, on peut écrire successivement :

$$H|\psi'_n\rangle \equiv H\mathcal{R}_i|\psi_n\rangle = \mathcal{R}_i H|\psi_n\rangle = \mathcal{R}_i E_n|\psi_n\rangle = E_n\mathcal{R}_i|\psi_n\rangle \equiv E_n|\psi'_n\rangle \ . \tag{17.203}$$

Cette chaîne d'égalités est vraie même si $\mathcal{R}_i$ est antiunitaire, puisque de toute façon E_n est une quantité réelle. En toute hypothèse, l'alternative est alors la suivante :

[74]Tout ce qui a été dit en considérant une seule fonction d'onde – c'est-à-dire un état à une seule composante – reste donc valide pour une particule de spin nul. Les complications liées au spin sont expliquées dans la section 20.4.

- ou bien $|\psi'_n\rangle$ coïncide en fait avec $|\psi_n\rangle$ et alors le niveau E_n n'est pas dégénéré. Ce cas doit être considéré comme relativement exceptionnel, surtout pour les problèmes dans plus d'une dimension d'espace[75] ;
- ou bien $|\psi'_n\rangle$ est différent de $|\psi_n\rangle$ et alors le niveau E_n présente une dégénérescence. Ceci est le cas le plus fréquent et implique l'existence de plusieurs états de même énergie ; pour distinguer ces états, on peut les noter $|\psi_{n,\,\alpha}\rangle$:

$$H|\psi_{n,\,\alpha}\rangle = E_n\,|\psi_{n,\,\alpha}\rangle \ . \tag{17.204}$$

Ce qui précède permet d'affirmer que toute opération de symétrie transforme les $|\psi_{n,\,\alpha}\rangle$ les uns dans les autres puisque, précisément, $\mathcal{R}_i|\psi_{n,\,\alpha}\rangle$ produit un vecteur propre d'énergie E_n, qui est la même que celle de $|\psi_{n,\,\alpha}\rangle$: une opération de symétrie laisse *globlement* invariant tout sous-espace propre. Autrement dit, on doit pouvoir écrire l'un d'entre eux en combinaison linéaire des autres :

$$\mathcal{R}_i|\psi_{n,\,\alpha}\rangle = \sum_\beta \mathcal{R}_{i,\,\beta\alpha}\,|\psi_{n,\,\beta}\rangle \ , \tag{17.205}$$

où les $\mathcal{R}_{i,\,\beta\alpha}$ sont des scalaires ; quels que soient ces derniers, la combinaison linéaire de droite est propre de H avec l'énergie E_n.

Cette observation est le socle de la Théorie de la représentation linéaire des groupes, très utile pour former d'emblée des vecteurs de symétrie adaptée, qu'il s'agisse de la symétrie de permutation de particules identiques (chapitre 22), ou des vibrations normales d'une molécule (voir problème 28.4.8) ; sur une telle base, la matrice de H est diagonale par blocs, ce qui facilite grandement la résolution effective de bien des problèmes. En effet, soit deux vecteurs propres $|\chi\rangle$ et $|\chi'\rangle$ d'une opération de symérie $\mathcal{R}$, associés à deux valeurs propres *différentes* :

$$\mathcal{R}|\chi\rangle = r|\chi\rangle \ , \qquad \mathcal{R}|\chi'\rangle = r'|\chi'\rangle \ , \qquad r \neq r' \ ; \tag{17.206}$$

comme $\mathcal{R}$ est une opération de symétrie, on a $[H,\,\mathcal{R}] = 0$. Prenons maintenant l'élément de matrice de cette dernière égalité entre $|\chi\rangle$ et $|\chi'\rangle$:

$$\langle\chi'|H\mathcal{R}|\chi\rangle - \langle\chi'|\mathcal{R}H|\chi\rangle = 0 \quad \Longleftrightarrow \quad (r-r')\langle\chi'|H|\chi\rangle = 0 \ . \tag{17.207}$$

Comme $r \neq r'$ par hypothèse, on en déduit $\langle\chi'|H|\chi\rangle = 0$: la matrice de H est bien diagonale par blocs vis-à-vis de deux sous-espaces engendrés par des vecteurs propres $\{|\chi_\alpha\rangle\}_\alpha$ et $\{|\chi'_\beta\rangle\}_\beta$, associés à deux valeurs propres r et r' différentes d'un même opérateur de symétrie.

17.7 Invariance $\mathcal{PT}$

La dernière section de ce chapitre consacré à la symétrie aborde des questions ayant donné lieu à une intense activité ces dernières années, à la suite des travaux pionniers de

[75]Dans une dimension d'espace, les états liés ne sont pas dégénérés (Tome I, chapitre 15, éq. (15.21)). L'argument ci-dessus a été utilisé pour montrer d'avance que les états propres de l'oscillateur linéaire ont forcément une parité déterminée (Tome I, chapitre 16, section 16.2).

Bender et Boettcher [18] sur les Hamiltoniens non hermitiques (pour une introduction très pédagogique, voir [19]).

L'un des "dogmes" de la théorie quantique est l'association d'un opérateur hermitique à toute grandeur physique, une propriété qui garantit la réalité des valeurs propres. En réalité, la condition d'hermiticité est une condition *suffisante*, nullement nécessaire, puisqu'il existe des opérateurs non hermitiques dont le spectre est réel. Ainsi, la matrice

$$\begin{bmatrix} 0 & -\mathrm{i} \\ 2\mathrm{i} & 0 \end{bmatrix} \tag{17.208}$$

qui n'est pas hermitique, a pour valeurs propres $\pm\sqrt{2}$. L'idée centrale est de remplacer la condition d'hermiticité par une condition plus faible assurant évidemment elle aussi la réalité des valeurs propres. L'un des premiers résultats importants fut obtenu dans l'étude d'une famille de Hamiltoniens de la forme :

$$H = \frac{p^2}{2m} - \lambda(\mathrm{i}x)^N \ , \tag{17.209}$$

où λ et N sont des paramètres positifs[76]. Un tel Hamiltonien est visiblement non-hermitique, puisque $H^\dagger = \frac{p^2}{2m} - \lambda(-\mathrm{i}x)^N \neq H$ si N n'est pas un entier pair. Par ailleurs, H n'est pas (en général) invariant par la parité Π, qui change p en $-p$ et x en $-x$ – sauf si $N = 4k+2$, $k \in \mathbb{N}$ –, ni par renversement du temps T qui change p en $-p$, laisse x inchangé mais complexe conjugue. En revanche, comme Bender et Boettcher [18] l'ont réalisé, tous les Hamiltoniens (17.209) sont invariants si on effectue les deux opérations Π *et* T, d'où la terminologie invariance $\mathcal{PT}$, pour reprendre leur notation. C'est ce qui les a conduits à proposer de remplacer la condition d'hermiticité par celle de l'invariance $\mathcal{PT}$, bien que, pour les Hamiltoniens de la forme (17.209), le spectre n'est réel que si N est supérieur à une certaine valeur $N_c \simeq 1,422\,07$, voir [19]. La non-suffisance de l'invariance $\mathcal{PT}$ pour la réalité du spectre force à distinguer les régions de l'espace des paramètres (pour (17.209), le seul paramètre est N) selon la réalité ou non du spectre, et récupère la notion de symétrie brisée, cette fois à propos de la symétrie $\mathcal{PT}$. Par la suite, d'autres Hamiltoniens ont été étudiés et fournissent des exemples où les régions de symétrie non-brisée sont aussi celles où le spectre est réel.

Afin de discuter plus précisément ces approches, et leur signification, prenons le cas $N = 4$, qui a été particulièrement étudié en raison de son caractère à première vue paradoxal. En effet, dans ce cas $H = \frac{p^2}{2m} - \lambda x^4$, $\lambda > 0$, et il saute aux yeux que, dans le cadre de travail ordinaire, un tel Hamiltonien ne possède aucun état lié. Physiquement, on peut même affirmer intuitivement que la question de la réalité du spectre ne se pose pas, le potentiel $V(x) \stackrel{\text{déf}}{=} -\lambda x^4$ donnant lieu aux phénomènes de diffusion ordinaires, avec un courant transmis et un courant réfléchi[77] ; en particulier, une particule d'énergie négative incidente émise par une source en $x = -\infty$ peut passer dans la région $x > 0$ par effet-tunel et y être détectée.

[76] Noter que le cas $N = 2$ correspond à l'oscillateur harmonique.

[77] On laise de côté les problèmes liés au fait que $V(x)$ n'est pas borné inférieurement.

En fait, la question de la réalité du spectre peut être posée dans un contexte plus général, en considérant toujours l'équation aux valeurs propres ordinaire $H\psi = E\psi$ mais en autorisant l'argument de la fonction ψ à être *complexe.* Cela étant admis, les conditions aux limites donnant (éventuellement) lieu à la quantification des valeurs propres s'énoncent en imposant à $\psi(z)$ de tendre vers zéro lorsque $|z| \rightarrow \infty$; la question est alors de savoir si ces valeurs discrètes sont complexes ou réelles. Dans le cas $N = 4$, ce spectre est effectivement réel, mais la condition aux limites n'est satisfaite que dans des domaines (secteurs) du plan complexe ne contenant pas l'axe réel. L'exclusion de l'axe réel permet d'accepter l'apparente contradiction entre ce que l'intuition exige – x réel, pas d'états liés, spectre continu et réel –, et les résultats établis – spectre discret réel (et d'ailleurs positif).

Ce cadre théorique élargi tente de définir une théorie quantique alternative, dont le contenu physique se doit d'être précisé[78] ; s'il est possible de prolonger une fonction d'onde à des valeurs complexes z de son argument – tout comme on effectue le prolongement analytique d'une fonction définie sur $\mathbb{R}$ –, le sens physique de celle-ci exige que z est réel, puisqu'il s'agit de définir une densité de probabilité de présence d'une particule autour d'un point dont la position ne peut être qu'un nombre réel.

Par ailleurs, il est certain que l'usage de Hamiltoniens non hermitiques ne date pas d'hier, notamment pour la description phénoménologique des systèmes instables, ou de la supraconductivité, mais l'usage qui en est alors fait n'est pas en opposition avec le contenu physique ordinaire de la théorie quantique. À l'heure actuelle, s'il semble prématuré de trancher définitivement sur la fécondité des idées fondées sur l'invariance $\mathcal{P}\mathcal{T}$, il convient d'ajouter que ces tentatives ont un intérêt théorique majeur, puisqu'elles explorent l'une des voies suggérées par Polyà et Hilbert pour démontrer la fameuse hypothèse de Riemann sur les zéros non-triviaux s_k de la fonction $\zeta(z)$ (pour un point récent, voir [20], [21]). La proposition de Polyà et Hilbert consiste à mettre en évidence un système physique dont les fréquences d'oscillation seraient les parties imaginaires des s_k ; dit autrement, il s'agit de trouver un Hamiltonien hermitique dont les valeurs propres seraient les nombres $-\mathrm{i}(s_k - \frac{1}{2})$. Parmi les tentatives les plus récentes, il convient de citer le travail de Castro qui construit une théorie quantique dite $\mathcal{C}\mathcal{T}$-invariante ($\mathcal{C}$ pour *charge*), présentée comme une généralisation de la version $\mathcal{P}\mathcal{T}$ de Bender (de nombreuses références sur ce sujet sont disponibles en [22]). Le problème 17.8.12 reprend quelques-unes des idées ci-dessus à propos d'un exemple très simple [23].

[78] *"It is not known whether non-Hermitian, $\mathcal{P}\mathcal{T}$- symmetric Hamiltonians can be used to describe experimentally observable phenomena"* ([19], p. 20).

17.8 Exercices et problèmes

17.8.1 Produits scalaire et vectoriel de deux opérateurs vectoriels

Montrer que le produit scalaire de deux opérateurs vectoriels est un opérateur scalaire au sens défini en (17.104), et que leur produit vectoriel est un opérateur vectoriel satisfaisant les relations caractéristiques (17.117).

17.8.2 Invariance de $[q,\, p] = \mathrm{i}\hbar 1$ par symétrie miroir

En utilisant (17.165), montrer que la relation fondamentale est invariante.

17.8.3 Opérateur de translation

On considère une particule se déplaçant suivant l'axe Ox ; soit $\mathbf{x}$ et $\mathbf{p}$ les opérateurs de position et d'impulsion. $|x\rangle$ désigne le ket propre de $\mathbf{x}$ associé à la valeur propre x. L'opérateur $\mathcal{T}(a)$ est défini par :

$$\mathcal{T}(a)|x\rangle = |x+a\rangle \ . \tag{17.210}$$

1. Montrer que $\mathcal{T}(a)\mathcal{T}(b) = \mathcal{T}(b)\mathcal{T}(a) = \mathcal{T}(a+b)$.

2. Donner l'expression des éléments de matrice $\langle x''|\mathcal{T}(a)|x'\rangle$ associés à $\mathcal{T}(a)$ en représentation-q. Montrer que $\mathcal{T}(a)^\dagger = \mathcal{T}(-a)$ et que $\mathcal{T}(a)$ est un opérateur unitaire.

3. Déterminer les éléments de matrice $\langle p''|\mathcal{T}(a)|p'\rangle$ associés à $\mathcal{T}(a)$ en représentation-p. En déduire que :
$$\mathcal{T}(a) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\mathbf{p}a} \ . \tag{17.211}$$

4. Soit $|\psi'\rangle = \mathcal{T}(a)|\psi\rangle$ le *ket* déduit de $|\psi\rangle$ par la transformation unitaire $\mathcal{T}(a)$; $\psi(x)$ et $\psi'(x)$ désignent les fonctions d'onde correspondantes. Montrer que :
$$\psi'(x) = \psi(x-a) \ . \tag{17.212}$$

5. Développer l'expression de $\mathcal{T}(a)$ donnée en 3 en série de puissances de $\mathbf{p}$. En utilisant l'expression de cet opérateur en représentation-x, déduire le développement de la fonction d'onde $\psi'(x)$ en série de puissances de a et retrouver le résultat de la question précédente.

17.8.4 Transformation de Galilée

Il s'agit de trouver la transformation G induite dans l'espace des états par une transformation de Galilée dans $\mathbb{R}^3$ faisant passer d'un repère R à un repère R' animé de la vitesse constante $\vec{v}_0$ par rapport à R. On pose précisément :

$$\Psi' = \mathcal{G}\,\Psi \ . \tag{17.213}$$

Ψ et Ψ' sont les fonctions d'onde, dans R et dans R' respectivement, d'une particule donnée (de masse m) animée d'un mouvement quelconque. Classiquement, la loi de transformation est :

$$\vec{r}\,' = \vec{r} - \vec{v}_0 t \ , \qquad \vec{p}\,' = \vec{p} - m\vec{v}_0 \ . \tag{17.214}$$

1. Écrire que les valeurs moyennes quantiques de la position et de l'impulsion se transforment comme les variables dynamiques classiques. En déduire les relations suivantes :
$$\mathcal{G}^\dagger \vec{r}\,\mathcal{G} = \vec{r} - \vec{v}_0 t \ , \qquad \mathcal{G}^\dagger \vec{p}\,\mathcal{G} = \vec{p} - m\vec{v}_0 \ . \tag{17.215}$$
2. On pose $\mathcal{G} = \mathrm{e}^{\mathrm{i}g}$. Expliquer pourquoi l'opérateur g doit être hermitique. Observer que g peut être choisi comme une forme linéaire de $\vec{r}$ et $\vec{p}$; déterminer g.
3. A l'aide de l'identité de Glauber[79], exprimer $\Psi'(\vec{r}, t)$ en fonction de $\Psi(\vec{r}, t)$.
4. Examiner en détail le cas d'une particule libre dont la fonction d'onde dans le repère R est une onde plane de vecteur d'onde $\vec{k}$.

17.8.5 Invariance de Galilée de l'équation de Schrödinger

Il s'agit ici d'établir l'invariance de l'équation de Schrödinger sous les transformations de Galilée. Par souci de simplicité on se place à une dimension d'espace. Soit un premier référentiel R dans lequel l'équation de Schrödinger pour une particule libre de masse m, s'écrit:

$$\mathrm{i}\hbar\frac{\partial\Psi(x,\,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,\,t)}{\partial x^2} \ . \tag{17.216}$$

On considère un second référentiel R' en translation uniforme à la vitesse V par rapport à R. Les coordonnées (x', t') utilisées dans R' sont liées aux coordonnées $(x,\,t)$ utilisées dans R par la transformation de Galilée :

$$x' = x - Vt \ , \qquad t' = t \ . \tag{17.217}$$

La loi de transformation correspondante des opérateurs différentiels est la suivante :

$$\frac{\partial}{\partial x'} = \frac{\partial}{\partial x} \ , \qquad \frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + V\frac{\partial}{\partial x} \ . \tag{17.218}$$

[79] Voir Tome I, éq. (13.215).

1. On pose que la nouvelle fonction d'onde $\Psi'(x', t')$ utilisée pour décrire la particule dans le référentiel R' est donnée par :

$$\Psi'(x', t') = \mathrm{e}^{\mathrm{i}f(x,t)}\ \Psi(x,\, t)\ , \tag{17.219}$$

où $f(x,\, t)$ est une fonction à valeurs réelles.

   (a) Justifier l'emploi d'une telle forme.
   (b) En imposant à $\Psi'(x', t')$ d'obéir à l'équation de Schrödinger pour une particule libre, écrire les conditions que doit satisfaire la fonction $f(x,\, t)$. On notera que les fonctions $\Psi(x,\, t)$ et $\partial\Psi(x,\, t)/\partial x$ sont *indépendantes.*
   (c) Déterminer explicitement la fonction $f(x,\, t)$.

2. L'équation de Schrödinger pour une particule libre admet des solutions du type onde plane :

$$\Psi(x,\, t) = \mathrm{e}^{\mathrm{i}(kx-\omega t)} \equiv \mathrm{e}^{2\mathrm{i}\pi\left(\frac{x}{\lambda}-\nu t\right)}\ , \tag{17.220}$$

où la longueur d'onde λ et la fréquence ν sont reliées à l'impulsion p et à l'énergie E de la particule via les relations de Planck - Einstein - de Broglie :

$$p = \frac{h}{\lambda}\ , \qquad E = h\nu\ . \tag{17.221}$$

C'est ce type de solution que l'on considère dans la suite.

   (a) Quelles sont les lois de transformation de l'impulsion p et de l'énergie E de la particule déduites *directement* du changement de référentiel galiléen (17.217) ?
   (b) La loi de transformation de $\Psi(x,\, t)$ lors d'un changement de référentiel galiléen est donc donnée par la relation (17.219) avec $f(x,\, t)$ obtenue ci-dessus. Ces résultats étant acquis, déterminer les lois de transformation de la longueur d'onde de de Broglie λ et de la fréquence de Planck - Einstein ν.
   (c) Que peut-on en déduire quant aux relations de Planck - Einstein - de Broglie lors d'une transformation de Galilée ?

17.8.6 Particule sur réseau unidimensionnel

Une particule de masse m se déplace dans un réseau périodique unidimensionnel (fig.17.1) de paramètre a.

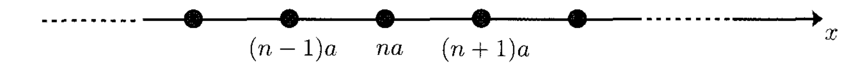


Figure 17.1: Réseau unidimensionnel de paramètre a

Le potentiel de réseau $V(x)$ étant périodique, on peut toujours l'écrire :

$$V(x) = \sum_{n\in\mathbb{Z}} v(x-na)\ . \tag{17.222}$$

1. Quel est le sens physique de la fonction $v(x)$?

2. Quels sont les opérateurs de translation $T(x_0)$ commutant avec le Hamiltonien ? Quelles sont leurs fonctions et valeurs propres ?

3. Compte tenu de ce qui précède, on cherche les fonctions propres communes à H (d'énergie E) et aux opérateurs de translation sous la forme :

$$\psi(x) = \mathrm{e}^{\mathrm{i}kx}\, u(x) \tag{17.223}$$

où $u(x)$ est une fonction à déterminer. À quelle condition sur u la fonction ψ est-elle fonction propre de $T(a)$? Quelle est la valeur propre associée à l'opérateur $T(a)$? Écrire l'équation différentielle satisfaite par $u(x)$.

4. Compte tenu de sa propriété établie ci-dessus, la fonction $u(x)$ peut être cherchée sous la forme d'une série de Fourier :

$$u(x) = \sum_{n\in\mathbb{Z}} c_n\, \mathrm{e}^{\mathrm{i}n(2\pi/a)\,x} \ , \tag{17.224}$$

où les $\{c_n\}_n$ sont des coefficients à déterminer. Écrire le système linéaire homogène satisfait par les $\{c_n\}_n$; on posera :

$$\mathcal{V}_r = \frac{1}{a}\int_{-a/2}^{+a/2} \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}r(2\pi/a)\,x}\, V(x) \ . \tag{17.225}$$

5. Dans toute la suite, on choisit[80] : $v(x) = g\,\delta(x)$, $(g > 0)$. $V(x)$ est donc un peigne de Dirac formé de barrières de potentiel équidistantes. Montrer que l'équation donnant les valeurs propres $E \equiv \hbar^2 K^2/(2m)$ est :

$$\frac{2mg}{a\hbar^2}\sum_{n\in\mathbb{Z}} \frac{1}{K^2 - [k + n\,(2\pi/a)]^2} = 1 \ . \tag{17.226}$$

6. Expliciter l'équation (17.226) sachant que[81] :

$$\sum_{n\in\mathbb{Z}} \frac{1}{(n+b)^2 - \lambda^2} = \frac{\pi}{2\lambda}\,[\cot\pi(b-\lambda) - \cot\pi(b+\lambda)] \ . \tag{17.227}$$

7. Discuter graphiquement les racines de cette équation. En déduire qu'il existe un ensemble de solutions paramétrées par k, notées $E_s(k)$; noter l'apparition d'intervalles en énergie où il n'existe pas de solutions ("bandes interdites").

8. Tracer qualitativement les E_s en fonction de k. Préciser les valeurs des E_s quand k est de la forme entier $\times$ (π/a). Comparer à la loi de dispersion d'une particule libre $E^{(0)} = \hbar^2 k^2/(2m)$.

9. Discuter brièvement le cas $g < 0$.

---

[80] Il s'agit alors précisément du modèle dit de Kronig - Penney.

[81] La somme de la série se trouve par résidus en calculant l'intégrale de la fonction $\frac{\pi\cot\pi z}{(z+b)^2-\lambda^2}$ sur le cercle à l'infini.

17.8.7 Particule sur réseau : une autre approche

Une particule de masse m se déplace sur un réseau régulier unidimensionnel. Dans la suite, $\phi(x)$, désigne une fonction bien localisée, ne prenant de valeurs non-nulles que sur un intervalle $\Delta x \ll a$ centré en $x = 0$.

On définit les fonctions $\phi_n(x) \stackrel{\text{déf}}{=} \phi(x - na)$, dont on suppose qu'elles constituent une base orthonormée complète de l'espace des états de la particule.

1. (a) Soit $\mathcal{T}(a)$ l'opérateur de translation d'amplitude a ; à quoi est égale la fonction $\mathcal{T}(a)\phi_n(x)$? En déduire la matrice de $\mathcal{T}(a)$ sur la base des $\phi_n(x)$.
   (b) τ désignant la valeur propre associée, on écrit les vecteurs propres de $\mathcal{T}(a)$ sous la forme :
   $$|\tau\rangle = \sum_n c_n|\phi_n\rangle \ . \tag{17.228}$$
   Quelle est la relation entre c_n et c_{n+1} ? Exprimer les c_n en fonction de τ et de c_0 et en déduire le développement de $|\tau\rangle$ sur les $|\phi_n\rangle$.
   (c) On pose désormais $\tau \equiv \tau_\alpha = \mathrm{e}^{-\mathrm{i}\alpha}$, où α est une variable continue. Expliquer pourquoi α est réel ; dans toute la suite, on considère exclusivement l'intervalle $]-\pi, +\pi[$ pour α.
   (d) Calculer le produit scalaire $(|\tau_\alpha\rangle, |\tau_{\alpha'}\rangle)$ et normaliser $|\tau_\alpha\rangle$ de sorte que[82] $(|\tau_\alpha\rangle, |\tau'_\alpha\rangle) = \delta(\alpha - \alpha')$.
   (e) Exprimer un $|\phi_n\rangle$ quelconque en fonction des $|\tau_\alpha\rangle$.
   (f) Sur la base $|\phi_n\rangle$, le Hamiltonien H de la particule est défini comme suit :
   $$H|\phi_n\rangle = \varepsilon|\phi_n\rangle - v(|\phi_{n+1}\rangle + |\phi_{n-1}\rangle) \ . \tag{17.229}$$
   où ε et v sont des grandeurs réelles positives ; v permet à la particule de sauter d'un site à ses premiers voisins. Exprimer H en fonction de $\mathcal{T}(a)$ et de son adjoint.
   (g) En déduire les valeurs propres $E(\alpha)$ et les vecteurs propres de H. Tracer la fonction $E(\alpha)$.
2. On suppose que la particule est initialement dans l'état $|\Psi(t=0)\rangle = |\phi_{n=0}\rangle$
   (a) Écrire le développement de l'opérateur d'évolution $U(t)$ sur les $|\tau_\alpha\rangle$.
   (b) En déduire le développement de $|\Psi(t)\rangle$ sur les $|\tau_\alpha\rangle$. Montrer que $|\Psi(t)\rangle$ peut finalement s'écrire :
   $$|\Psi(t)\rangle = \mathrm{e}^{-\mathrm{i}\theta(t)} \sum_{n\in\mathbb{Z}} C_n(t)|\phi_n\rangle \ , \tag{17.230}$$
   où les coefficients $C_n(t)$ s'expriment simplement à l'aide des fonctions de Bessel ordinaires[83] $J_n(\gamma t)$; préciser la phase $\theta(t)$ et la constante γ.

[82] On rappelle que $\sum_{n\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}nx} = 2\pi\delta(x)$ quand $-\pi \le x \le +\pi$.
[83] On peut les définir comme $J_n(T) \stackrel{\text{déf}}{=} \frac{1}{2\pi}\int_{-\pi}^{+\pi} \mathrm{e}^{-\mathrm{i}nx}\, \mathrm{e}^{\mathrm{i}T\sin x}\, \mathrm{d}x$.

(c) Donner l'expression de la probabilité $P_0(t)$ de trouver, à l'instant t, la particule localisée sur son site de départ.

(d) Soit la fonction $G(u\,;t)$ définie comme suit :

$$G(u\,;t) \stackrel{\text{déf}}{=} \sum_{n\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}nu} J_n(\gamma t) \;; \tag{17.231}$$

montrer que $G(u\,;t) = \mathrm{e}^{\mathrm{i}\gamma t\,\sin u}$. En déduire, pour vérification, que $|\Psi(t)\rangle$ est bien normé à tout temps.

(e) Combien vaut la valeur moyenne de la coordonnée $\langle x\rangle(t)$ dans l'état $|\Psi(t)\rangle$?

(f) Exprimer $\langle x^2\rangle(t)$ sous la forme d'une série des fonctions de Bessel (on utilisera $\langle\phi_n|x^2|\phi_n\rangle \simeq n^2a^2$).

(g) En utilisant $\frac{\partial G}{\partial u}$, montrer que l'écart quadratique $\Delta x^2(t)$ est égal à At^2, où A est une constante à préciser.

17.8.8 Renversement du temps

1. Soit un système dont le Hamiltonien H est invariant par renversement du temps et soit $\vec{p}$ une observable représentée sur la base propre de H par une matrice dont tous les éléments sont imaginaires purs. Montrer que dans tout état stationnaire non-dégénéré de H, la valeur moyenne de $\vec{p}$ est nulle.

2. Montrer que, en représentation-p, l'opérateur K de renversement du temps agit comme suit :

$$K\,\Phi(\vec{p}, t) = \Phi^*(-\vec{p}, -t) \;. \tag{17.232}$$

17.8.9 Dynamique d'un électron dans une cage atomique

Un agrégat d'atomes possède la géométrie schématisée sur la figure 17.2 : un atome central, repéré par l'indice 0, est entouré de N atomes répartis régulièrement sur un cercle, numérotés de 1 à N. L'angle entre deux atomes périphériques consécutifs est $\theta = 2\pi/N$. Chaque atome n ($n = 0, 1, \ldots, N$) possède un seul état, représenté par une fonction bien localisée autour de ce site. Le ket correspondant est noté $|n\rangle$; ces $N+1$ vecteurs sont supposés orthonormalisés :

$$\langle n|n'\rangle = \delta_{nn'} \qquad (n,\, n' = 0,\, 1,\, 2,\, \ldots,\, N) \;. \tag{17.233}$$

Un électron[84], lâché sur cette structure, peut se déplacer du site central aux sites périphériques et inversement, l'amplitude de probabilité correspondante étant $\hbar\omega$; le

[84]On ignore le degré de liberté de spin, supposant que tous les processus envisagés ne changent pas le spin ; celui-ci étant alors une constante du mouvement, il est légitime de ne pas l'introduire explicitement dans le traitement.

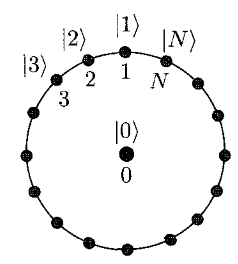


Figure 17.2: Configuration géométrique de l'agrégat.

Hamiltonien de l'électron est donc :

$$H = \hbar\omega \sum_{n=1}^{N} (|0\rangle\langle n| + |n\rangle\langle 0|) \ . \tag{17.234}$$

Il s'agit d'étudier le mouvement de l'électron, un état initial $|\Psi(t=0)\rangle$ étant donné.

1. On désigne par $\mathcal{R}$ l'opérateur de rotation dans l'espace des états associé à la rotation de θ (dans le sens trigonométrique direct) autour d'un axe perpendiculaire au plan de l'agrégat et passant par 0. Compléter les équations suivantes :

   $$\mathcal{R}|0\rangle = \ldots\ldots \ , \qquad \mathcal{R}|n\rangle = \ldots\ldots \qquad (n = 1, 2, \ldots N) \ . \tag{17.235}$$

   En déduire la matrice $M_{\mathcal{R}}$ représentant $\mathcal{R}$ sur la base $\{|n\rangle\}_{n=0,\,1,\,\ldots,\,N}$.

2. Expliquer pourquoi H et $\mathcal{R}$ commutent entre eux. Si une explication claire et précise suffit, montrer néanmoins par un calcul explicite que $[H, \mathcal{R}] = 0$.

3. Montrer que les vecteurs suivants :

   $$|\psi_0\rangle = |0\rangle \ , \qquad |\psi_k\rangle = C \sum_{n=1}^{N} \mathrm{e}^{\mathrm{i}kn\theta}\,|n\rangle \qquad (k = 1, 2, \ldots, N) \tag{17.236}$$

   sont les vecteurs propres de $\mathcal{R}$; préciser les valeurs propres r_k, $k = 0, 1, 2, \ldots, N$. Combien vaut la constante de normalisation C ? Quelle est la dégénérescence de chaque valeur propre ?

4. En partant du commutateur $[H, \mathcal{R}]$, montrer que le vecteur $H|\psi_k\rangle$ est propre de $\mathcal{R}$. En déduire que pour $k = 1, 2, \ldots, N-1$, $|\psi_k\rangle$ est propre de H. Quelle est l'énergie propre associée à ce $|\psi_k\rangle$?

5. Les deux derniers états propres de H sont donc des combinaisons linéaires de $|\psi_0\rangle$ et de $|\psi_N\rangle$. Trouver ces combinaisons linéaires (les noter $|\psi_\pm\rangle$)[85], ainsi que les valeurs propres associées, $\varepsilon_\pm$.

[85] ne pas oublier de les normaliser.

6. Compléter l'expression suivante de l'opérateur d'évolution :

$$U(t) = \sum_{\alpha=\pm} |\psi_\alpha\rangle\langle\psi_\alpha|\,\mathrm{e}^{\cdots} + \sum_{k=1}^{N-1} |\psi_k\rangle\langle\psi_k|\,\mathrm{e}^{\cdots} \ . \tag{17.237}$$

7. Initialement, l'électron est localisé sur l'atome central : $|\Psi(t=0)\rangle = |0\rangle$. Donner le développement de $|\Psi(t)\rangle$ sur les états $|\psi_\pm\rangle$, $|\psi_k\rangle$ ($k = 1, 2, \ldots, N-1$). En déduire la probabilité $P(t)$ de trouver, à l'instant t, l'électron sur son atome de départ.

8. Dans le cas où l'électron part d'un site quelconque de la périphérie, calculer la probabilité $p(t)$ de le trouver à l'instant t sur l'atome central.

17.8.10 Groupe des rotations planes

Ce problème formalise quelques résultats obtenus dans le problème 17.8.9.

Soit un polygone régulier à N sommets numérotés de 0 à $N-1$ dans le sens positif. Dans toute la suite, R désigne la rotation dans $\mathbb{R}^2$ de l'angle $\frac{2\pi}{N}$ autour de l'axe perpendiculaire au plan du polygone, et passant par le centre de celui-ci. Sur chaque sommet n, on définit une fonction $\psi_n(\vec{r}) \equiv \langle\vec{r}\,|\psi_n\rangle$ obtenue par rotation à partir de la même fonction définie en $n = 0$ par la rotation $\mathcal{R}^n$. Les $|\psi_n\rangle$ sont orthonormalisés.

1. Combien y a-t-il de rotations distinctes ? Montrer que toutes s'obtiennent à partir de l'une d'entre elles.

2. On définit naturellement le produit de deux rotations comme $\mathcal{R}^n\mathcal{R}^m \stackrel{\text{déf}}{=} \mathcal{R}^{n+m}$. Montrer que l'ensemble $\mathcal{G}$ des rotations a alors la structure de groupe. Est-il commutatif ?

3. Citer une caractéristique des valeurs propres des rotations.

4. Utiliser le théorème de Cayley - Hamilton pour trouver le spectre $\{r_k\}_k$ des rotations.

5. On forme les opérateurs $P_k \stackrel{\text{déf}}{=} C\sum_{n=0}^{N-1} r_k^n\mathcal{R}^n$; trouver la constante C de sorte que $P_k^2 = P_k$.

6. Que vaut le produit $P_k P_k'$?

7. Soit les vecteurs $|\Psi_k\rangle \stackrel{\text{déf}}{=} \mathcal{N}P_k|\psi_0\rangle$; montrer que chacun d'entre eux est propre de toutes les rotations, et préciser à chaque fois la valeur propre associée. Que vaut $\mathcal{N}$ quand $|\Psi_k\rangle$ est de norme unité ?

8. Soit $\mathcal{S}_N$ le groupe des permutations de N objets. Montrer[86] que $\mathcal{G}$ est un sous-groupe de $\mathcal{S}_N$.

[86] Ce résultat est connu sous le nom de *théorème de Cayley*.

9. Soit ϕ l'angle azimutal dans le plan mesuré à partir du sommet $n = 0$, et on note désormais $\psi(\vec{r}, \phi_n)$ la fonction $\psi_n(\vec{r})$. Par définition (voir (17.93)), on a $\mathcal{R}(\delta\phi)\psi(\vec{r}, \phi_n) = \psi(\vec{r}, \phi_n - \delta\phi)$, avec $\delta\phi = \frac{2\pi}{N}$; en déduire que, dans la limite $N \to +\infty$, la rotation d'un angle fini ϕ est l'opérateur $\mathcal{R}(\phi) = \mathrm{e}^{-\phi\frac{\mathrm{d}}{\mathrm{d}\phi}} \equiv \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\phi L_z}$.

17.8.11 Dilatations

On considère deux observateurs O et O' liés aux repères R (coordonnées x, p) et R' (coordonnées x', p') immobiles l'un par rapport à l'autre. Ces deux observateurs étudient un seul et même système S mais utilisent des unités de longueur et d'impulsion différentes. Ces unités sont telles que les *valeurs numériques* prises par les coordonnées classiques (x, p) et (x', p') des deux observateurs sont reliés par la loi de transformation :

$$x' = \lambda x \ , \qquad p' = \frac{p}{\lambda} \ . \tag{17.238}$$

où λ est un paramètre réel. La transformation (17.238) est appelée *dilatation.*

Les observateurs O et O' décrivent l'état quantique du système S au moyen des *mêmes* opérateurs x et p et par des kets différents notés $|\psi\rangle$ et $|\psi'\rangle$ respectivement. Les kets $|\psi\rangle$ et $|\psi'\rangle$ sont reliés par la transformation supposée unitaire $\mathcal{D}_\lambda$:

$$|\psi'\rangle = \mathcal{D}_\lambda|\psi\rangle \ . \tag{17.239}$$

Dans la suite, il s'agit de trouver explicitement $\mathcal{D}_\lambda$ et d'étudier ses propriétés.

1. En termes de fonctions d'onde, la transformation s'écrit :

$$\psi'(x) \stackrel{\text{déf}}{=} \mathcal{D}_\lambda\psi(x) = C_\lambda\psi(\lambda^{-1}x) \ ; \tag{17.240}$$

   en utilisant la normalisation des fonctions d'onde, trouver la constante C_λ.

2. L'invariance des valeurs moyennes se traduit par $\langle x'\rangle' = \langle x\rangle$, soit :

$$\langle\psi'|x'|\psi'\rangle = \langle\psi|x|\psi\rangle \ ;$$

   en déduire que :

$$\mathcal{D}_\lambda^\dagger x\mathcal{D}_\lambda = \lambda x \ , \qquad \mathcal{D}_\lambda^\dagger p\mathcal{D}_\lambda = \frac{p}{\lambda} \ . \tag{17.241}$$

3. Montrer que $\mathcal{D}_\lambda D_{\lambda'} = D_{\lambda\lambda'}$. On choisit désormais de paramétrer les dilatations au moyen de la quantité α telle que $\lambda = \mathrm{e}^\alpha$, $\alpha \in \mathbb{R}_+$ et on pose $\mathcal{D}_\lambda = \Delta_\alpha$. A quelle valeur de α correspond l'identité ? Que vaut le produit $\Delta_\alpha\Delta_{\alpha'}$?

4. On pose :

$$\Delta_\alpha = \mathrm{e}^{\frac{\alpha}{\mathrm{i}\hbar}d} \ , \tag{17.242}$$

   d étant un opérateur hermitique.

(a) En utilisant l'expression (17.242) et les relations (17.241), récrites avec Δ_α et développées en série entière de α, établir les relations de commutation $[d, x]$ et $[d, p]$.

(b) Compte tenu de ce qui précède, montrer que le générateur d peut être cherché sous la forme :

$$d = (1 - \nu)\, xp + \nu\, px \tag{17.243}$$

(c) Déterminer la valeur du paramètre ν, supposé réel, pour que d soit hermitique.

5. Il s'agit maintenant de retrouver autrement l'action de $\mathcal{D}_\lambda$ sur une fonction $\psi(x)$ quelconque.

(a) Montrer que l'opérateur Δ_α peut s'écrire :

$$\Delta_\alpha = \mathrm{e}^{-\frac{\alpha}{2}}\, \mathrm{e}^{\frac{\alpha}{i\hbar}xp} \equiv \mathrm{e}^{-\frac{\alpha}{2}}\, \mathrm{e}^{-\alpha\Omega} \tag{17.244}$$

Quelle est, en représentation-q, l'expression de l'opérateur Ω ?

(b) Pour élucider l'action de Δ_α sur une fonction quelconque $\psi(x)$, on pose :

$$\mathrm{e}^{-\alpha\Omega}\, \psi(x) = \psi(f(\alpha,\, x)) \qquad \forall\, \psi \tag{17.245}$$

où $f(\alpha,\, x)$ est une fonction à déterminer. Combien vaut[87] $f(0,\, x)$?

(c) En effectuant une dérivation partielle en α, montrer que l'équation (17.245) est équivalente à :

$$x\,\frac{\partial f}{\partial x}\,\psi' \;=\; -\frac{\partial f}{\partial \alpha}\,\psi' \qquad \forall \psi' \tag{17.246}$$

où ψ' désigne ici la dérivée de la fonction ψ.

(d) Vérifier que :

$$f(\alpha,\, x) \;=\; x\,\mathrm{e}^{-\alpha} \tag{17.247}$$

est la solution de (17.246) satisfaisant la condition aux limites obtenue en 5b.

(e) Retrouver l'action de $\mathcal{D}_\lambda$ sur une fonction quelconque obtenue en 1 :

$$\mathcal{D}_\lambda \psi(x) \;=\; \frac{1}{\sqrt{\lambda}}\,\psi\left(\frac{x}{\lambda}\right) \; . \tag{17.248}$$

6. *Application*

Une particule de masse m est plongée dans un potentiel unidimensionnel $V(x)$; son Hamiltonien est noté H.

(a) Calculer le commutateur $[d, H]$.

(b) Prenant sa valeur moyenne dans un état propre du Hamiltonien, montrer que pour tout état stationnaire $|\psi\rangle$ pour lequel $\langle\psi|d|\psi\rangle$ existe, on a la relation suivante entre valeurs moyennes[88] (elle conduit au *Théorème du Viriel*) :

$$\langle \frac{p^2}{2m} \rangle = \langle x V'(x) \rangle \; . \tag{17.249}$$

[87] Ceci définit la condition aux limites pour la fonction inconnue f.

[88] Cette égalité a été obtenue autrement dans le Tome I (voir éq. (14.164)).

(c) Appliquant cette relation à un potentiel $V(x) = \frac{1}{2}kx^2$, montrer que dans ce cas les valeurs moyennes de l'énergie cinétique et de l'énergie potentielle sont égales entre elles et donc chacune égale à la moitié de l'énergie propre de l'état. Généraliser aux potentiels de la forme $V(x) = C|x|^n$.

17.8.12 Un exemple à propos de l'invariance $\mathcal{PT}$

Il s'agit d'illustrer les questions évoquées dans la section 17.7 à propos de l'exemple le plus simple qui soit [23].

1. Soit un Hamiltonien dont la représentation sur une base donnée est la matrice :

$$H \overset{\text{déf}}{=} \begin{bmatrix} r\mathrm{e}^{\mathrm{i}\theta} & s \\ s & r\mathrm{e}^{-\mathrm{i}\theta} \end{bmatrix} , \tag{17.250}$$

où les trois paramètres r, s et θ sont réels (on pourra supposer $s > 0$). Sachant que, sur la même base, l'opérateur parité $\mathcal{P}$ a la matrice :

$$\mathcal{P} \overset{\text{déf}}{=} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} , \tag{17.251}$$

montrer que H est $\mathcal{PT}$ symétrique.

2. Trouver les valeurs propres de H.

3. En déduire l'existence de deux régions dans l'espace des paramètres, l'une où la symétrie $\mathcal{PT}$ est brisée, l'autre où elle ne l'est pas.

4. Trouver les vecteurs propres de H, $|\varepsilon_\pm\rangle$, dans cette dernière région (on posera $\sin\alpha \overset{\text{déf}}{=} \frac{r}{s}\sin\theta$). À quelle condition sont-ils propres de l'opérateur $\mathcal{PT}$ pour la valeur propre $+1$?

5. Ces vecteurs propres sont-ils orthogonaux vis-à-vis du produit scalaire habituel ?

6. On définit un produit scalaire $[|u\rangle, |v\rangle] \overset{\text{déf}}{=} (\mathcal{PT}|u\rangle).|v\rangle$ où $|f\rangle.|g\rangle$ désigne la somme des produits deux à deux des composantes. Montrer que les $|\varepsilon_\pm\rangle$ sont orthogonaux au sens de ce produit scalaire. Calculer le carré de la norme de chacun d'entre eux (observer que la norme ainsi définie n'est plus forcément positive), et ajuster les constantes de proportionnalité de sorte que les normes soient de module unité.

7. Soit l'opérateur ("de charge") $\mathcal{C}$:

$$\mathcal{C} \overset{\text{déf}}{=} \frac{1}{\cos\alpha} \begin{bmatrix} \mathrm{i}\sin\alpha & 1 \\ 1 & -\mathrm{i}\sin\alpha \end{bmatrix} . \tag{17.252}$$

Montrer que H et $\mathcal{C}$ commutent. Combien vaut $\mathcal{C}^2$? Quels sont les valeurs et vecteurs propres de $\mathcal{C}$?

8. Soit un autre produit scalaire défini comme $\langle u|v\rangle \stackrel{\text{déf}}{=} \{|u\rangle, |v\rangle\} \stackrel{\text{déf}}{=} \mathcal{CPT}|u\rangle.|v\rangle$. Montrer que maintenant $\{|\varepsilon_\pm\rangle, |\varepsilon_\pm\rangle\} = 1$; démontrer *explicitement* que quel que soit $|\psi\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$, $\langle\psi|\psi\rangle \geq 0$.

9. En déduire que vis-à-vis du produit scalaire $\mathcal{CPT}$, la relation de fermeture a maintenant la forme familière :

$$|\varepsilon_+\rangle\langle\varepsilon_+| + |\varepsilon_-\rangle\langle\varepsilon_-| = \mathbf{1} \ . \tag{17.253}$$

physique

Chapitre 18

Théorie du moment cinétique

Ce chapitre est consacré à la théorie générale du moment cinétique en Mécanique quantique. Les états propres d'un moment cinétique orbital (harmoniques sphériques) sont introduits avant d'établir le théorème fondamental d'addition. On démontre enfin l'important théorème de Wigner - Eckart qui est l'un des fondements des règles de sélection.

18.1 Importance du moment cinétique

On connaît l'importance du moment cinétique $\vec{L} \stackrel{\text{déf}}{=} \vec{r} \times m\vec{v}$ en Mécanique classique, eu égard au théorème dit du moment cinétique :

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\vec{L} = \vec{r} \times \vec{F}} \tag{18.1}$$

qui est une conséquence immédiate de l'équation fondamentale de la dynamique (17.2). En particulier, pour une force centrale, associée à une énergie potentielle V ne dépendant que de la *distance* au centre (pas de l'orientation du vecteur $\vec{r}$), $V(r)$, la seule composante non nulle de la force $\vec{F} = -\vec{\nabla}V(r)$ est colinéaire à $\vec{r}$ et le produit vectoriel $\vec{r} \times \vec{F}$ est nul : le *vecteur* $\vec{L}$ est alors une constante du mouvement. C'est pourquoi la trajectoire est plane, située dans un plan perpendiculaire à la valeur constante de $\vec{L}$, fixée une fois pour toutes par les conditions initiales ; c'est aussi la raison pour laquelle la trajectoire obéit à la loi des aires : la surface Σ balayée par le rayon-vecteur $\vec{r}$ entre deux instants t_0 et t est proportionnelle au temps ($\Sigma = \|\vec{L}\|(t-t_0)/m$, voir Tome I, problème 7.4.9 pour le champ coulombien). $\vec{L}$ est l'équivalent pour des variables angulaires sur la sphère du moment conjugué $\vec{p} = (p_x,\, p_y,\, p_z)$ pour des coordonnées cartésiennes.

Si le potentiel est seulement à symétrie cylindrique autour d'un axe choisi comme axe Oz, la force est dirigée suivant cet axe et (18.1) dit que la dérivée du vecteur $\vec{L}$ est perpendiculaire à Oz : dans ces conditions, seule la composante L_z est constante dans le temps, autre exemple montrant que la réduction de la symétrie s'accompagne inéluctablement d'une diminution du nombre des constantes du mouvement.

Bien évidemment, le moment cinétique joue également un rôle fondamental en Mécanique quantique, et ce d'autant plus que des expériences cruciales ont démontré sa quantification (par exemple : expériences de Stern et Gerlach, voir Tome I, chapitre 11, section 11.2), tout comme l'énergie des états liés d'un système. En outre, c'est le moment cinétique (orbital ou de spin) qui permet de comprendre la nature des sources du magnétisme (voir chapitre 27, section 27.7), ce que la Physique classique est à nouveau dans l'incapacité de décrire, en raison du théorème de Miss van Leeuwen. S'agissant du magnétisme, et comme on l'a vu au chapitre précédent, les statuts des grandeurs classiques et celui des observables sont inversés en ce qui concerne l'invariance de jauge. Dans le cadre quantique, afin de manipuler fondamentalement des grandeurs invariantes de jauge, on appellera moment cinétique (orbital) l'observable[1] $\vec{L} \stackrel{\text{déf}}{=} \vec{r} \times \vec{p}$. Sur le plan technique, la quantification de $\vec{L}$ vient du fait que les variables angulaires sont *confinées* dans un intervalle fini (par exemple, l'angle azimutal $\phi \in [0, 2\pi]$), tout comme l'est le nombre d'onde k pour une particule qui, étant strictement localisée dans un puits infini, a une abscisse restreinte à l'intervalle $[0, a]$ (Tome I, sous-section 15.3.2).

Ce point étant précisé, on a montré ci-dessus (voir notamment (17.98)) que l'opérateur ainsi défini est le générateur des rotations spatiales. L'importance de la symétrie, en Mécanique quantique comme ailleurs, n'étant plus à démontrer, on pressent le rôle central que jouera le moment cinétique dès qu'il s'agira d'exploiter ce type de symétrie.

Enfin, force est effectivement d'admettre (au vu notamment de certaines complexités des spectres atomiques) que les particules possèdent un degré de liberté intrinsèque, en tout point analogue à un moment cinétique, mais sans équivalent classique ; pour des raisons historiques, ce degré de liberté a été appelé *spin*[2] et est noté traditionnellement $\vec{S}$. C'est d'ailleurs ce degré de liberté qui conduit à définir deux grandes classes de particules, les fermions et les bosons, dont les comportements, radicalement différents, produisent des effets spectaculaires[3]. Pour rappeler sa nature en terme de grandeurs dy-

[1]Bien que les opérateurs $\vec{r}$ et $\vec{p}$ ne commutent pas, il n'est pas nécessaire de symétriser l'observable $\vec{L}$ en écrivant $\vec{L} = \frac{1}{2}(\vec{r} \times \vec{p} - \vec{p} \times \vec{r})$. En effet, chaque composante de $\vec{L}$ est une combinaison linéaire de produits $r_u p_v$ avec à chaque fois $u \neq v$.

[2]À l'époque où cette terminologie a été choisie, on a imaginé que ce moment cinétique représentait une rotation de la particule sur elle-même, ce que les astronomes désignent par *spin*. En réalité, le spin d'une particule n'est susceptible d'aucune représentation de cette sorte (voir par exemple le problème 11.3.2 du Tome I).

Comme on le verra par la suite (voir notamment la Remarque 1 p. 866), c'est au fond parce que la Physique classique ($\hbar = 0$) peut se passer des nombres complexes que le spin n'y apparaît pas.

[3]C'est parce que les électrons sont des fermions que les métaux sont conducteurs de l'électricité (voir chapitre 29). C'est aussi parce qu'ils peuvent se condenser par paires que certains métaux deviennent supraconducteurs à (très) basse température. Enfin, c'est la coopération subtile de ces deux classes de particules qui permet de comprendre la stabilité de la matière sous toutes ses formes, y compris les plus exotiques – en particulier, c'est elle qui assure l'existence de la limite thermodynamique quand des charges de signes contraires sont en jeu [24].

namiques, $\vec{L}$ sera plus précisément appelé moment cinétique *orbital*, rappelant le fait qu'il est l'analogue quantique d'un moment cinétique d'une particule qui, classiquement, serait en orbite autour d'un centre fixe.

$\vec{L}$ est un opérateur ayant trois composantes L_x, L_y et L_z ($L_x = yp_z - zp_y$, etc) dont la définition à partir de son expression classique ne pose aucun problème puisque dans chaque produit, les deux opérateurs commutent. Il en résulte immédiatement que ces composantes sont, telles quelles, des opérateurs hermitiques. En utilisant la relation fondamentale $[u, p_v] = \mathrm{i}\hbar\,\delta_{uv}\mathbf{1}$, il est facile de montrer[4] que L_x, L_y et L_z satisfont les relations de commutation suivantes, qui se déduisent les unes des autres par permutations circulaires :

$$[L_x,\, L_y] = \mathrm{i}\hbar\, L_z \ , \qquad [L_y,\, L_z] = \mathrm{i}\hbar\, L_x \ , \qquad [L_z,\, L_x] = \mathrm{i}\hbar\, L_y \ . \tag{18.2}$$

Le $\hbar$ au second membre est bien nécessaire pour l'homogénéité ; quant au i, il ne doit pas surprendre : tous les opérateurs sont hermitiques, mais le commutateur de deux opérateurs hermitiques est antihermitique[5].

Par contraste, on désignera habituellement par $\vec{S}$ un moment cinétique de spin, aussi appelé *moment cinétique intrinsèque*, pour rappeler notamment qu'il ne peut être considéré comme le résultat d'une combinaison vectorielle de variables dynamiques, au contraire de $\vec{L}$. Pour des raisons physiques, il est nécessaire d'admettre que les trois composantes de $\vec{S}$ satisfont les mêmes relations de commutation que celles de $\vec{L}$ (voir chapitre 20) : c'est pourquoi il est possible de faire une théorie générale du moment cinétique en Mécanique quantique, quelle que soit l'origine microscopique de celui-ci ; quand la nature (orbitale ou de spin) est non-pertinente, l'usage est de désigner par $\vec{J}$ un tel moment cinétique abstrait, dont les trois composantes obéissent aux relations *caractéristiques* :

$$\boxed{[J_x,\, J_y] = \mathrm{i}\hbar\, J_z \quad [J_y,\, J_z] = \mathrm{i}\hbar\, J_x \quad [J_z,\, J_x] = \mathrm{i}\hbar\, J_y \quad \Longleftrightarrow \quad \vec{J}\times\vec{J} = \mathrm{i}\hbar\vec{J}} \tag{18.3}$$

Rappelons au passage que $\vec{J}$ est un opérateur vectoriel (voir p. 683, et plus particulièrement éq. (17.117) où le rôle de $\vec{L}$ est ici joué par $\vec{J}$).

En introduisant le tenseur antisymétrique[6] ε_{uvw} qui vaut 0 si deux indices (ou plus) sont égaux, +1 si uvw se déduit de xyz par une permutation circulaire et -1 dans le cas contraire, on a :

$$\boxed{[J_u,\, J_v] = \mathrm{i}\hbar\,\varepsilon_{uvw}\, J_w} \tag{18.4}$$

En utilisant maintenant $[J_u,\, J_v^2] = [J_u,\, J_v]J_v + J_v[J_u,\, J_v]$, on voit facilement que :

$$\boxed{[J_u,\, \vec{J}^2] = 0 \quad (u = x,\, y,\, z) \quad \Longleftrightarrow \quad [\vec{J}, \vec{J}^2] = 0} \tag{18.5}$$

---

[4]Voir l'exemple avec $\vec{J}$, traité dans la sous-section 12.1.2 du Tome I.
[5]au sens où $\Omega^\dagger = -\,\Omega$.
[6]Voir aussi le problème 7.4.6 du Tome I, et son corrigé.

où $\vec{J}^2 \stackrel{\text{déf}}{=} J_x^2 + J_y^2 + J_z^2$. Il existe donc des vecteurs propres communs à $\vec{J}^2$ et à l'*une* de ses autres composantes ; l'habitude est de choisir le couple $(\vec{J}^2, J_z)$, mais il est clair que ceci est purement conventionnel. En outre, il se révèle très commode d'introduire les deux combinaisons remarquables :

$$\boxed{J_+ \stackrel{\text{déf}}{=} J_x + \mathrm{i}\, J_y \qquad J_- \stackrel{\text{déf}}{=} J_x - \mathrm{i}\, J_y} \tag{18.6}$$

J_+ et J_- sont hermitiques conjugués l'un de l'autre, mais leur produit est hermitique :

$$J_+^\dagger = J_- \ , \qquad (J_-J_+)^\dagger = J_-J_+ \ , \qquad (J_+J_-)^\dagger = J_+J_- \ . \tag{18.7}$$

Les $J_\pm$ satisfont des relations déduites facilement de (18.3) :

$$\boxed{[J_z,\, J_\pm] = \pm\hbar\, J_\pm \qquad [J_+,\, J_-] = 2\hbar\, J_z \qquad [J_\pm,\, \vec{J}^2] = 0} \tag{18.8}$$

On a aussi $J_x^2 + J_y^2 = \frac{1}{2}(J_+J_- + J_-J_+)$, d'où l'expression très utile en pratique :

$$\boxed{\vec{J}^2 = J_z^2 + \frac{1}{2}\,(J_+J_- + J_-J_+)} \tag{18.9}$$

Avec $J_+J_- = J_-J_+ + 2\hbar\, J_z$ et (18.9), on trouve aussi :

$$\vec{J}^2 = J_z^2 + \hbar J_z + J_-J_+ = J_z^2 - \hbar J_z + J_+J_- \quad \Longleftrightarrow \quad J_\pm J_\mp = \vec{J}^2 - J_z^2 \pm \hbar J_z \tag{18.10}$$

Clairement, les trois composantes J_x, J_y et J_z ont le même spectre (isotropie de l'espace), et d'ailleurs sont équivalents les uns aux autres au sens où il existe un opérateur (unitaire) tel que $J_u = \mathcal{R}_{uv} J_v \mathcal{R}_{uv}^\dagger$. Ainsi, on a par exemple $J_y = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\frac{\pi}{2}J_z} J_x\, \mathrm{e}^{-\frac{1}{\mathrm{i}\hbar}\frac{\pi}{2}J_z}$, en conformité avec (17.98) puisque J_y se déduit de J_x par une rotation de $\frac{\pi}{2}$ autour de Oz. Cette égalité peut être vérifiée en utilisant le développement (13.204) du problème 13.9.3 (Tome I), et en prenant en compte les relations caractéristiques (18.3).

Les relations de commutation permettent aussi d'écrire des inégalités portant sur les écarts quadratiques des composantes du moment cinétique. Avec la mesure habituelle des fluctuations $\Delta A \stackrel{\text{déf}}{=} \sqrt{\langle A^2\rangle - \langle A\rangle^2}$, où les $\langle \ldots \rangle$ désignent la moyenne avec un état quelconque, et compte tenu d'un théorème général (voir Tome I, éq. (10.96)), on a :

$$\boxed{\Delta J_x\, \Delta J_y \geq \frac{1}{2}\, |\langle J_z\rangle|} \tag{18.11}$$

et les inégalités analogues déduites par permutation des indices.

Comme les J_u ne commutent pas entre eux, il n'existe pas de vecteurs propres communs à ces trois observables. Il existe toutefois des états exceptionnels $|\psi\rangle$ où les écarts quadratiques des *trois* composantes sont simultanément nuls. D'après (18.11), une condition nécessaire pour ceci est $\langle \vec{J}\rangle = 0$; pour avoir également $\Delta\vec{J} = 0$, il faut de surcroît $\langle J_x^2\rangle = \langle J_y^2\rangle = \langle J_z^2\rangle = 0$. Pour un tel état $|\psi\rangle$, on a :

$$\vec{J}\,|\psi\rangle = 0 \ . \tag{18.12}$$

Un tel état est dit *sphérique*, et se comporte à la manière d'un scalaire dans toute rotation, puisqu'il est invariant :

$$\mathcal{R}_{\vec{u},\,\theta}|\psi\rangle \equiv \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\vec{u}.\vec{J}}|\psi\rangle = |\psi\rangle \qquad \forall\vec{u}\ ,\forall\theta\ . \tag{18.13}$$

Enfin, une autre relation remarquable est l'inégalité :

$$\langle\vec{J}^2\rangle \geq \langle J_u^2\rangle \geq \langle J_u\rangle^2\ , \tag{18.14}$$

qui résulte de $\vec{J}^2 - J_z^2 = J_x^2 + J_y^2$; l'opérateur au second membre étant positif, le premier l'est tout autant et la valeur moyenne d'un opérateur positif est un nombre positif. Dans (18.14), l'égalité n'est obtenue que si les deux autres composantes ont un carré moyen nul, ce que donne l'état exceptionnel $|\psi\rangle$ (18.12). Dans tous les autres cas, on a :

$$\langle\vec{J}^2\rangle > \langle J_u\rangle^2 \quad\Longleftrightarrow\quad \sqrt{\langle\vec{J}^2\rangle} > |\langle J_z\rangle|\ , \tag{18.15}$$

et les relations analogues avec J_x et J_y : la racine carrée de la valeur moyenne du carré du module d'un moment cinétique est donc toujours plus grande que la valeur moyenne de n'importe laquelle de ses composantes. La moyenne $\langle\vec{J}\,\rangle$ peut être colinéaire à l'un des axes de coordonnées, par exemple $\langle J_z\rangle \neq 0$, $\langle J_x\rangle = \langle J_y\rangle = 0$, mais dans tous les cas $\langle\|\vec{J}^2\|\rangle^{1/2} > |\langle J_z\rangle|$. Le fait est que, à l'instar de variables aléatoires, le carré de la valeur moyenne n'est pas égal à la valeur moyenne du carré ; dans l'exemple précédent, on peut se figurer que J_x et J_y varient ("fluctuent"), tout en ayant une moyenne nulle[7].

18.2 Propriétés générales des valeurs et vecteurs propres d'un moment cinétique

On va montrer, de façon purement algébrique[8], en quoi consiste le spectre d'un moment cinétique, c'est-à-dire quelles sont les valeurs possibles que l'on peut trouver lors d'une mesure de l'observable correspondante (module ou composante selon un axe donné). Ce qui va être établi vaut autant pour le moment orbital que pour le moment de spin : seules les relations de commutation caractéristiques (18.3) vont être utilisées, avec, sous-jacente, l'hypothèse que les vecteurs propres sont tous des vecteurs de norme *finie* – et comme d'habitude fixée conventionnellement à l'unité. La méthode utilisée est une méthode d'échelle (ou d'escalier) en bien des points similaire à celle qui a été utilisée pour la

[7]On verra qu'en présence d'un champ magnétique parallèle à Oz, le vecteur $\langle\vec{J}\rangle$ précesse autour du champ en faisant un angle constant avec celui-ci : les composantes transverses ont chacune une moyenne nulle (elles varient harmoniquement en temps à une fréquence proportionnelle au champ) et la composante longitudinale est bien, en effet, inférieure (en module) à la longueur du vecteur $\langle\vec{J}\rangle$ (voir aussi la Remarque p. 725). Noter que si les moyennes $\langle J_x\rangle = \langle J_y\rangle = 0$ sont nulles, ce n'est pas parce qu'il s'agit de moyennes *temporelles* !

[8]par opposition à une méthode purement analytique pour le moment orbital, conduisant aux harmoniques sphériques Y_{lm} (voir section 18.3).

quantification de l'oscillateur harmonique selon Dirac, introduisant naturellement les opérateurs de création et d'annihilation a et $a^\dagger$ (Tome I, section 16.3).

Comme $\vec{J}^2$ et J_z commutent, ces deux opérateurs ont des vecteurs propres communs ; désignons ces derniers par $|\lambda\, m\rangle$: λ est un nombre pur qui donne la valeur propre associée à l'opérateur $\vec{J}^2$, m donne de même celle associée à J_z ; pour des raisons purement dimensionnelles, on a :

$$\boxed{\vec{J}^2\,|\lambda\, m\rangle = \lambda\hbar^2\,|\lambda\, m\rangle \qquad J_z\,|\lambda\, m\rangle = m\hbar\,|\lambda\, m\rangle} \tag{18.16}$$

Les nombres λ et m sont réels (les opérateurs sont hermitiques) mais à part ceci quelconques, pour l'instant ; toutefois, en conséquence de l'inégalité (18.14), ils satisfont :

$$\lambda \geq m^2 \ , \tag{18.17}$$

ce qui montre (sans surprise) que λ est positif ou nul et que $|m| \leq \sqrt{\lambda}$. Par ailleurs, le sens positif le long de l'axe Oz est purement conventionnel : la symétrie galiléenne exige donc que, pour λ fixé, si la valeur m est possible, la valeur $-m$ est également possible.

Mettons maintenant en œuvre la procédure d'échelle annoncée ([4], p. 375 et sq.) ; partons de la relation $J_zJ_+ = J_+J_z + \hbar J_+$ (voir (18.8)) appliquée à $|\lambda\, m\rangle$; compte tenu de la deuxième relation (18.16), elle donne :

$$J_zJ_+|\lambda\, m\rangle = (J_+J_z + \hbar J_+)|\lambda\, m\rangle = (m+1)\hbar\, J_+|\lambda\, m\rangle \ . \tag{18.18}$$

De la même façon, en partant de $J_zJ_- = J_-J_z - \hbar J_-$, on obtient :

$$J_zJ_-|\lambda\, m\rangle = (m-1)\hbar\, J_-|\lambda\, m\rangle \ . \tag{18.19}$$

Enfin, avec $[\vec{J}^2,\, J_\pm] = 0$, on trouve :

$$\vec{J}^2 J_\pm|\lambda\, m\rangle = \lambda\hbar^2\, J_\pm|\lambda\, m\rangle \ . \tag{18.20}$$

Ainsi, $|\lambda\, m\rangle$ étant un vecteur propre de $(\vec{J}^2,\, J_z)$ avec les valeurs propres $\lambda\hbar^2$ et $m\hbar$, $J_\pm|\lambda\, m\rangle$ est également un vecteur propre de ces deux opérateurs, mais associé aux valeurs propres $\lambda\hbar^2$ et $(m \pm 1)\hbar$. Les deux relations suivantes sont donc vraies :

$$\boxed{J_\pm\,|\lambda\, m\rangle = C_\pm(\lambda, m)\hbar\,|\lambda, m \pm 1\rangle} \tag{18.21}$$

où les $C_\pm(\lambda, m)$ sont des scalaires, fonctions de λ et m à déterminer.

Précisons maintenant le lien entre λ et m, dont l'existence est indubitable pour des raisons physiques. Compte tenu de (18.17), la valeur absolue de m est bornée quand λ est donné. Désignons par j la valeur maximale de m et par $|\lambda, j\rangle$ le vecteur propre correspondant – qui n'est pas le vecteur nul ; à ce stade on sait que (voir (18.17)) :

$$m_{\text{max}} \stackrel{\text{déf}}{=} j \leq \sqrt{\lambda} \ . \tag{18.22}$$

Comme J_+ augmente le nombre m d'une unité, l'application de J_+ à ce dernier vecteur ne peut produire un nouveau vecteur propre ; on doit donc avoir :

$$J_+ \, |\lambda \, j\rangle = 0 \ . \tag{18.23}$$

Appliquons maintenant J_- à cette équation et prenons en compte (18.10) :

$$J_- J_+ \, |\lambda \, j\rangle = 0 \quad \Longleftrightarrow \quad (\vec{J}^2 - J_z^2 - \hbar J_z) \, |\lambda \, j\rangle = 0 \ , \tag{18.24}$$

soit :

$$(\lambda - j^2 - j)\hbar^2 \, |\lambda, j\rangle = 0 \quad \Longleftrightarrow \quad \lambda = j(j+1) \tag{18.25}$$

Ceci donne la relation précise entre la valeur maximale de m, j, et le nombre λ donnant la valeur propre de $\vec{J}^2$; j étant positif ou nul, l'inégalité (18.22) est visiblement satisfaite.

De la même façon, soit m_{min} la plus petite valeur prise par m ; notons-la j' et par définition $-\sqrt{\lambda} \leq j' \leq j$; une argumentation analogue montre que :

$$J_- \, |\lambda \, j'\rangle = 0 \quad \Longrightarrow \quad J_+ J_- \, |\lambda \, j'\rangle = 0 \quad \Longleftrightarrow \quad (\vec{J}^2 - J_z^2 + \hbar J_z) \, |\lambda \, j'\rangle = 0 \ , \tag{18.26}$$

soit :

$$(\lambda - j'^2 + j')\hbar^2 \, |\lambda \, j'\rangle = 0 \quad \Longleftrightarrow \quad \lambda = j'^2 - j' = j'(j'-1) \ , \tag{18.27}$$

ce qui fournit une relation précise entre j' et λ. Les deux expressions (18.25) et (18.27) de λ ne sont compatibles que si $j(j+1) = j'(j'-1)$, soit $j' = -j$ ou $j' = j+1$. Cette dernière possibilité est à exclure, car elle dit que j' est plus grand que j ; il reste donc $j' = -j$, soit $m_{\text{min}} = -j \equiv -m_{\text{max}}$, en conformité d'ailleurs avec l'invariance galiléenne, d'où les résultats très importants :

$$\boxed{\lambda = j(j+1) \qquad -j \leq m \leq j} \tag{18.28}$$

Il est ainsi possible d'atteindre $|\lambda, -j\rangle$ – le vecteur propre ayant la plus petite valeur de m – à partir de $|\lambda \, j\rangle$ en appliquant l'opérateur J_- un nombre suffisant de fois. Comme à chaque action de J_-, m décroît d'une unité, la différence $m_{\text{max}} - m_{\text{min}}$, égale à $j - (-j)$, est nécessairement un entier positif ou nul, d'où le résultat majeur : j est soit *entier*, soit *demi-entier* :

$$\boxed{j = 0, \, 1, \, 2, \, \ldots \quad \Longleftrightarrow \quad \lambda \equiv j(j+1) = 0, \, 2, \, 6, \, \ldots} \tag{18.29}$$

ou

$$\boxed{j = \frac{1}{2}, \, \frac{3}{2}, \, \frac{5}{2}, \, \ldots \quad \Longleftrightarrow \quad \lambda \equiv j(j+1) = \frac{3}{4}, \, \frac{15}{4}, \, \frac{30}{4}, \, \ldots} \ . \tag{18.30}$$

Par ailleurs, comme deux valeurs m et $-m$ doivent être simultanément présentes, les valeurs de m sont réparties symétriquement autour de zéro ; au total, pour un j donné, les valeurs propres possibles de J_z sont en nombre égal à $2j+1$ et données par :

$$j\hbar, \, (j-1)\hbar, \, (j-2)\hbar, \, \ldots, \, (-j+1)\hbar, \, -j\hbar \ . \tag{18.31}$$

Ainsi, le nombre quantique[9] m prend des valeurs entières si j est entier, demi-entières si j est demi-entier :

$$\boxed{m = j,\, j-1,\, j-2,\, \ldots,\, -j+1,\, -j} \tag{18.32}$$

À j fixé (entier *ou* demi-entier), m varie strictement d'une unité quand on passe d'un vecteur propre au suivant et peut prendre au total $g_j \stackrel{\text{déf}}{=} 2j+1$ valeurs distinctes.

Trouvons maintenant les scalaires $C_\pm$ introduits en (18.21) ; le carré scalaire du vecteur au second membre de cette équation est :

$$|C_+(\lambda\, m)|^2\, \hbar^2\, (|m+1\rangle,\, |m+1\rangle) = (J_+|\lambda\, m\rangle,\, J_+|\lambda\, m\rangle) = \langle \lambda, m|J_- J_+|\lambda\, m\rangle \ . \tag{18.33}$$

En supposant tous les vecteurs propres implicitement normalisés à l'unité[10], et en utilisant $J_+^\dagger = J_-$, il vient :

$$|C_+(\lambda, m)|^2\, \hbar^2 = \langle \lambda, m|J_- J_+|\lambda\, m\rangle = [j(j+1) - m^2 - m]\hbar^2 \ , \tag{18.34}$$

la dernière égalité venant de (18.10). De ceci, on ne peut pas déduire la phase de C_+, qui est sans importance ; par usage, on adopte le choix le plus simple en posant :

$$C_+(\lambda, m) = \sqrt{j(j+1) - m(m+1)} \ ; \tag{18.35}$$

par des moyens analogues[11], on trouve :

$$C_-(\lambda, m) = \sqrt{j(j+1) - m(m-1)} \ . \tag{18.36}$$

Maintenant que la relation entre λ et j est précisée, changeons légèrement les notations afin de se conformer à l'usage et notons désormais $|j\, m\rangle$ les états propres $|\lambda\, m\rangle$ (bien que $\lambda \neq j$, si $j \neq 0$) :

$$\boxed{J_z|j\, m\rangle = m\hbar|j\, m\rangle \qquad \vec{J}^2|j\, m\rangle = j(j+1)\hbar^2|j\, m\rangle} \tag{18.37}$$

L'usage désigne souvent par *multiplet* l'ensemble des $2j+1$ vecteurs $\{j,\, m\rangle\}_{-j\leq m\leq +j}$, mais le même terme désigne aussi parfois l'ensemble des composantes de structure fine issu d'une approximation purement électrostatique (voir chapitre 27, section 27.3). Noter que dans la limite quasi-classique ($j \gg 1$), la différence entre $j(j+1)$ et j^2 s'estompe ; elle est en revanche essentielle pour $j \sim 1$, puisqu'elle résulte directement des relations de commutation caractéristiques (lesquelles sont nulles si $\hbar = 0$). Comme indiqué dans le Tome I, cette distinction est l'une des empreintes de la théorie quantique, reconnue pour la première fois par Heisenberg pour le couple $(q,\, p)$ (voir chapitre 9, section 9.2).

Dans ces nouvelles notations, les relations fondamentales *à retenir* sont :

$$\boxed{J_\pm|j\, m\rangle = \hbar\, \sqrt{j(j+1) - m(m\pm 1)}\ |j\, m\pm 1\rangle} \tag{18.38}$$

[9]appelé nombre quantique magnétique pour des raisons qui seront claires par la suite.

[10]C'est à ce point précis que l'on a besoin de l'hypothèse suivant laquelle ces vecteurs sont de norme finie.

[11]ou en remarquant que par symétrie galiléenne, on doit avoir $C_-(\lambda, m) = C_+(\lambda, -m)$.

Elles permettent d'engendrer les $2j+1$ vecteurs propres de $(\vec{J}^2,\, J_z)$, pour un j donné, à partir de l'un d'entre eux, par exemple $|j\, m_{\text{max}} = j\rangle$ à qui l'on applique J_- autant de fois que nécessaire. Notons que ces dernières égalités s'écrivent aussi :

$$J_\pm |j\, m\rangle = \hbar \sqrt{(j \mp m)(j \pm m + 1)}\; |j\, m \pm 1\rangle \ , \tag{18.39}$$

ou encore :

$$J_\pm |j\, m\rangle = \hbar \sqrt{\vec{J}^2 - J_z(J_z \pm \hbar \mathbf{1})}\; |j\, m \pm 1\rangle = \sqrt{J_x^2 + J_y^2 \mp \hbar J_z}\; |j\, m \pm 1\rangle \ , \tag{18.40}$$

en appliquant la règle habituelle : si $|\omega\rangle$ est un vecteur propre de l'opérateur Ω avec la valeur propre ω, alors $f(\Omega)|\omega\rangle = f(\omega)|\omega\rangle$, f étant une fonction quelconque.

De ces diverses relations, il sera fait un usage intensif dans la suite et dans les applications. Sur la base $\{|j\, m\rangle\}_{-j \leq m \leq +j}$ ordonnée par valeurs décroissantes de m, les matrices représentant J_+ et J_- sont purement non-diagonales et n'ont d'éléments non-nuls que juste au-dessus et au-dessous (respectivement) de la diagonale principale ; la première (resp. dernière) colonne de la matrice de J_+ (resp. J_-) est nulle.

Il est clair que, à j fixé, le multiplet des $2j+1$ vecteurs $\{|j\, m\rangle\}_{-j \leq m \leq +j}$ forme un sous-espace dégénéré pour la valeur propre $\hbar^2 j(j+1)$ de $\vec{J}^2$. Toujours à j fixé, on voit que ce sous-espace est stable par l'application de J_x et J_y, qui sont des combinaisons linéaires des $J_\pm$, puisque $J_\pm|j\, m\rangle$ est un vecteur du *même* sous-espace (les opérateurs $J_\pm$ font glisser m mais ne changent pas j). En l'absence d'autres degrés de liberté (associés à d'autres observables pertinentes), $\vec{J}^2$ et J_z forment un ECOC[12], puisque la donnée des nombres quantiques j et m spécifie complètement leur état propre commun : il n'y a pas de dégénérescence et on ne peut ajouter une autre composante $J_{u \neq z}$.

♦ *Remarques*

1. La relation (18.15) prend maintenant la forme explicite :

$$\sqrt{\langle \vec{J}^2 \rangle} = \sqrt{j(j+1)}\, \hbar > j\hbar = \hbar \max_{-j \leq m \leq +j} m = \max \langle J_z \rangle \ . \tag{18.41}$$

   Le rapport $\langle J_z \rangle / \langle \vec{J}^2 \rangle^{1/2}$, qui évoque le cosinus de l'angle entre le moment cinétique et sa projection sur Oz, est donc toujours strictement plus petit que 1 : sa plus grande valeur est $\sqrt{j/(j+1)}$ (elle tend vers 1 à la limite classique $j \gg 1$).

2. J_x et J_y ont toujours une moyenne nulle dans un état $|j\, m\rangle$: ce sont des combinaisons linéaires des opérateurs $J_\pm$ qui changent m d'une unité ; en conséquence, $\langle j,\, m | J_\pm | j\, m\rangle = 0$ quels que soient j et m, et il en va de même pour tous les éléments diagonaux de J_x et J_y sur la base propre de J_z. On a donc :

$$\boxed{\langle \vec{J} \rangle = \langle J_z \rangle\, \vec{e}_z = m\hbar\, \vec{e}_z} \tag{18.42}$$

[12] Suivant la définition de [25], II. D, p. 144.

où les moyennes sont calculées dans un état $|j\,m\rangle$ quelconque et où $\vec{e}_z$ est le vecteur unitaire porté par l'axe de quantification. Le petit dessin souvent rencontré (un vecteur $\vec{J}$ avec une orientation quantifiée) peut être trompeur, sauf à imaginer que c'est un instantané d'un vecteur $\vec{J}$ tournant à toute vitesse autour de Oz.

En revanche, les écarts-types de J_x et J_y sont *finis* (sauf pour les états sphériques définis en (18.12)). Dans un état $|j\,m\rangle$, $\langle J_x\rangle = 0$, donc $\Delta J_x^2 \stackrel{\text{déf}}{=} \langle J_x^2\rangle - \langle J_x\rangle^2 = \langle J_x^2\rangle$, et de même pour ΔJ_y^2. Les deux directions Ox et Oy sont orthogonales mais arbitraires vis-à-vis de l'axe de quantification, de sorte que $\langle J_x^2\rangle = \langle J_y^2\rangle$. Utilisant $J_x^2 + J_y^2 = \vec{J}^2 - J_z^2$ et prenant la valeur moyenne membre à membre dans $|j\,m\rangle$, il vient :

$$\langle J_x^2\rangle = \langle J_y^2\rangle = \frac{1}{2}\langle \vec{J}^2 - J_z^2\rangle = \frac{1}{2}\hbar^2[j(j+1) - m^2] \qquad (18.43)$$

d'où le résultat :

$$\Delta J_x = \Delta J_y = \frac{\hbar}{\sqrt{2}}\sqrt{j(j+1) - m^2} \ ; \qquad (18.44)$$

les fluctuations de J_x et J_y sont donc minimum si $|m| = j$, maximum si $m = 0$. Comme $j \geq |m| \equiv \hbar^{-1}|\langle J_z\rangle|$, on en déduit :

$$\boxed{\Delta J_x \Delta J_y \geq \frac{\hbar}{2}\,|\langle J_z\rangle|} \qquad (18.45)$$

en accord avec l'inégalité (10.97) démontrée dans le Tome I, vraie d'ailleurs quel que soit l'état avec lequel on prend les valeurs moyennes.

3. Revenant à l'expression (17.123) de l'opérateur d'une rotation de 2π, on voit que si m est entier, alors $\mathcal{R}_{\vec{u},\,2\pi} = \mathbf{1}$, et si m est demi-entier, $\mathcal{R}_{\vec{u},\,2\pi} = -\mathbf{1}$. Pour un système donné, pour une rotation de 2π, c'est *l'ensemble* des états qui peut au plus acquérir une phase unique ; ceci confirme que c'est bien l'ensemble des valeurs de m qui est composé soit d'entiers, soit de demi-entiers. Le mélange de ces deux types de nombres n'est pas possible et constitue une *règle de supersélection* interdisant toute communication entre des espaces d'événements associés chacun au caractère entier *ou* demi-entier de j. Physiquement, ceci signifie, par exemple, que partant d'un ensemble de particules toutes de moment cinétique entier (*bosons*), la théorie quantique standard interdit tout processus conduisant à un ensemble de spins demi-entiers (*fermions*) en nombre impair. Il existe toutefois des scénarios permettant l'émergence de fermions (spins demi-entiers) dans une théorie ne contenant que des champs de bosons [26]. ♦

18.3 États propres du moment cinétique orbital

Le moment cinétique orbital, noté $\vec{L}$, est un moment cinétique particulier. On peut donc d'ores et déjà affirmer que ses états propres, notés $|l\,m\rangle$ selon l'usage, satisfont :

$$\boxed{\vec{L}^2|l\,m\rangle = l(l+1)\hbar^2\,|l\,m\rangle \qquad L_z|l\,m\rangle = m\hbar\,|l\,m\rangle} \qquad (18.46)$$

où, selon (18.30), l est un entier (ou un demi-entier, pour l'instant on n'en sait pas plus) positif ou nul, et m un entier (ou un demi-entier) assujetti aux inégalités $-l \leq m \leq +l$, prenant ainsi $2l+1$ valeurs distinctes. Toutefois, on va montrer que, pour un moment cinétique *orbital* – qui est égal à un produit $\vec{r} \times \vec{p}$ –, seules les valeurs *entières* de l (et donc de m) sont possibles.

$\vec{L}^2$ et L_z sont deux opérateurs dont on connaît l'expression en représentation-q, par exemple ; il est donc facile d'écrire les équations différentielles exprimant les équations aux valeurs propres, préliminaire à leur résolution d'une façon ou d'une autre.

Comme rappelé au début de ce chapitre, le moment cinétique orbital joue un rôle fondamental dans tous les problèmes à force centrale, qui possèdent de ce fait une symétrie sphérique. En pareil cas, les coordonnées adaptées au problème sont évidemment les coordonnées sphériques, définies comme d'habitude par r, θ et ϕ ($r \geq 0$, $0 \leq \phi \leq 2\pi$, $0 \leq \theta \leq \pi$) :

$$\boxed{x = r\sin\theta\cos\phi \qquad y = r\sin\theta\sin\phi \qquad z = r\cos\theta} \tag{18.47}$$

À r constant, le point repéré par les angles θ et ϕ se déplace sur la sphère de rayon r ; par référence au globe terrestre, les lignes $\phi = \mathrm{C}^{\text{ste}}$ correspondent aux méridiens, l'origine $\phi = 0$ (méridien de Greenwich) étant choisie dans le plan xOz. L'angle θ est la latitude : $\theta = 0$ correspond au pôle Nord, $\theta = \pi/2$ à l'équateur, $\theta = \pi$ au pôle Sud ; les lignes iso-θ sont les parallèles. L'élément de volume $\mathrm{d}^3r = \mathrm{d}x\mathrm{d}y\mathrm{d}z$ est égal à :

$$\boxed{\mathrm{d}^3r = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi} \tag{18.48}$$

Le vecteur $\vec{L}$ est un opérateur différentiel, tout comme $\vec{p}$, mais il n'agit que sur les angles θ et ϕ, ce qui est *a priori* évident quand on se souvient que $\vec{L}$ entre dans l'expression des opérateurs de rotation, et qu'aucune rotation ne modifie la distance r au centre : $\vec{L}$ ne saurait donc altérer r.

On connaît au départ les expressions des composantes cartésiennes de $\vec{L}$, que l'on sait exprimer à l'aide des coordonnées cartésiennes et de leurs moments conjugués. Par exemple :

$$L_x = yp_z - zp_y = -\mathrm{i}\hbar\left(y\,\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \tag{18.49}$$

puisque, en représentation-q, $p_u = -\mathrm{i}\hbar\partial_u$, et de même pour L_y et L_z. En procédant comme d'habitude quand on passe d'un jeu de variables à un autre, il est facile, mais un peu pénible, d'obtenir L_x, L_y et L_z exprimés en coordonnées sphériques (voir exercice 18.7.1) ; on trouve[13] :

$$L_x = \mathrm{i}\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cos\phi\cot\theta\frac{\partial}{\partial\phi}\right)\ , \qquad L_y = \mathrm{i}\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \sin\phi\cot\theta\frac{\partial}{\partial\phi}\right)\ , \tag{18.50}$$

$$L_z = -\mathrm{i}\hbar\,\frac{\partial}{\partial\phi}\ . \tag{18.51}$$

[13]En représentation-p, $x = +\mathrm{i}\hbar\partial_{p_x}$, ... ; si le vecteur $\vec{p}$ est repéré par les angles θ et ϕ de la même façon, toutes les expressions des L_{p_u} changent de signe.

L'expression de L_z ne doit pas surprendre, car elle peut en fait s'obtenir très vite en raisonnant comme suit. On sait du chapitre 17 (voir (17.98)) que l'opérateur de la rotation d'un angle ϕ autour de Oz est $e^{\frac{1}{i\hbar}\phi L_z}$; pour une rotation infinitésimale d'angle $\delta\phi$, cet opérateur est $\mathbf{1} + \frac{\delta\phi}{i\hbar} L_z + \mathcal{O}(\delta\phi^2)$. Par ailleurs, selon (17.93), la transformée de toute fonction $f(\phi)$ par cette rotation est $f(\phi - \delta\phi)$, soit $f(\phi) - \delta\phi\, f'(\phi) + \ldots$. D'où l'égalité :

$$\big(\mathbf{1} + \frac{\delta\phi}{i\hbar} L_z + \ldots\big) f(\phi) = f(\phi) - \delta\phi\, f'(\phi) + \ldots \ , \tag{18.52}$$

puis, par identification, l'expression (18.51) de L_z. On note que, dans les expressions (18.50) et (18.51), le coefficient de chaque opérateur différentiel ∂_θ, ∂_ϕ et imaginaire pur : tout comme pour $p_u = -i\hbar\partial_u$, cette association assure l'hermiticité des opérateurs représentant les grandeurs physiques (observables) dans les espaces de fonctions douées des bonnes propriétés aux limites[14] :

$$\Big(i\frac{\partial}{\partial q}\Big)^\dagger = i\frac{\partial}{\partial q} \quad \Longleftrightarrow \quad \Big(\frac{\partial}{\partial q}\Big)^\dagger = -\frac{\partial}{\partial q} \tag{18.53}$$

le signe $-$ à droite se combinant à celui de l'intégration par parties pour satisfaire l'égalité traduisant la condition d'hermiticité.

♦ *Remarque*

Comme l'ont montré Judge et Lewis [27] et Lévy-Leblond [28], la manipulation de variables angulaires exige de prendre quelques précautions afin d'éviter des contradictions (pour plus de détails, voir problème 18.7.2). Pour la simplicité, considérons seulement l'angle azimutal ϕ et l'opérateur $L_z = -i\hbar\partial_\phi$, qui est le moment conjugué p_ϕ de ϕ, en convenant une fois pour toutes de prendre ϕ dans l'intervalle $[0,\, 2\pi]$.

Une première observation est que L_z n'est hermitique que dans un espace de fonctions 2π-périodiques[15], $f(\phi+2\pi) = f(\phi)$; en effet, avec deux fonctions quelconques $f(\phi)$ et $g(\phi)$, la propriété d'hermiticité de L_z se traduit par :

$$\langle f | L_z g\rangle = \langle L_z f | g\rangle \Longleftrightarrow \int_0^{2\pi} f^*(\phi)\Big(-i\hbar\frac{\partial g}{\partial\phi}\Big)\mathrm{d}\phi = \int_0^{2\pi}\Big(-i\hbar\frac{\partial f}{\partial\phi}\Big)^* g(\phi)\mathrm{d}\phi \ . \tag{18.54}$$

En effectuant une intégration par parties, on voit que ceci est vrai ssi le produit $f(\phi)g(\phi)$ reprend la même valeur en 0 et en 2π, quel que soit le couple $(f,\, g)$. En prenant pour f la fonction égale à 1 partout, on voit cette condition est satisfaite ssi g est 2π-périodique : $L_z = -i\hbar\partial_\phi$ n'est hermitique que dans un espace exclusivement composé de fonctions 2π-périodiques. Il en résulte que, en représentation-q, l'opérateur angle azimutal $\hat{\phi}$ est associé à la multiplication par l'angle ϕ *restreint* à un intervalle de mesure 2π, $[0,\, 2\pi]$ par exemple ; c'est donc la fonction périodique en dents de scie valant ϕ sur $]0,\, 2\pi[$ et prolongée par périodes sur $\mathbb{R}$:

$$\hat{\phi} \longrightarrow \times\phi_{\mathrm{per}}(\phi)\ :\ \phi_{\mathrm{per}}(\phi) \stackrel{\text{déf}}{=} \phi,\ 0 < \phi < 2\pi,\ \phi_{\mathrm{per}}(\phi + 2\pi) = \phi_{\mathrm{per}}(\phi) \ , \tag{18.55}$$

[14] On rencontrera dans la suite d'autres exemples montrant l'importance des conditions aux limites pour les fonctions sur lesquelles agissent les opérateurs (voir Remarque ci-après et problème 18.7.2).

[15] Tout comme, en représentation-q, $p_x = -i\hbar\partial_x$ n'est hermitique que pour des fonctions s'annulant aux bornes du domaine accessible.

une définition qui traduit le confinement de la variable angulaire. Un point important à noter est que la fonction $\phi_{\text{per}}(\phi)$ est *discontinue*. Cela étant précisé, le produit des opérateurs ∂_ϕ et $\hat{\phi}$ agit comme suit sur une fonction $f(\phi)$:

$$\frac{\partial}{\partial\phi}\hat{\phi}f(\phi) = \Big(1 - 2\pi\sum_{n\in\mathbb{Z}}\delta(\phi - 2\pi n)\Big)f(\phi) + \hat{\phi}f'(\phi) \ . \tag{18.56}$$

La grande parenthèse en facteur de $f(\phi)$ est la dérivée de la fonction périodique $\phi_{\text{per}}(\phi)$ représentant l'angle, définie ci-dessus. Par ailleurs, $\hat{\phi}L_z f\phi) = -\mathrm{i}\hbar\hat{\phi}f'(\phi)$, d'où le commutateur $[\hat{\phi},\, L_z]$:

$$\boxed{[\hat{\phi},\, L_z] = -\mathrm{i}\hbar\Big(1 - 2\pi\sum_{n\in\mathbb{Z}}\delta(\phi - 2\pi n)\Big)\,\mathbf{1}} \tag{18.57}$$

La périodicité de la variable angulaire se traduit donc par une relation de commutation compliquée avec son moment conjugué – l'erreur à ne pas commettre étant d'oublier le peigne de Dirac en écrivant hâtivement $[\hat{\phi},\, L_z] = -\mathrm{i}\hbar\mathbf{1}$. Noter que quand ϕ n'apparaît que dans des fonctions 2π-périodiques, par exemple $\mathrm{e}^{\mathrm{i}n\phi}$, $n \in \mathbb{Z}$, toutes ces complications disparaissent de fait. ♦

Par ailleurs, il est facile de vérifier que les expressions (18.50) et (18.51) satisfont les relations de commutation (18.2) – voir exercice 18.7.1 ; elles permettent aussi d'obtenir[16] :

$$\boxed{\vec{L}^2 = -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta\,\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)} \tag{18.58}$$

$$\boxed{L_\pm = \pm\hbar\,\mathrm{e}^{\pm\mathrm{i}\phi}\Big(\frac{\partial}{\partial\theta} \pm \mathrm{i}\,\cot\theta\,\frac{\partial}{\partial\phi}\Big)} \tag{18.59}$$

En outre, et toujours par des intégrations par parties, on vérifie que dans l'espace des fonctions 2π-périodiques, les L_u sont hermitiques et que $(L_\pm)^\dagger = L_\mp$, comme il se doit (voir aussi (18.53) et exercice 18.7.1).

Ces expressions étant obtenues, il est possible d'écrire explicitement les équations aux valeurs et fonctions propres pour $\vec{L}^2$ et L_z, qui commutent entre eux. Leurs fonctions propres communes, indicées par les nombres quantiques l et m, associées aux valeurs propres selon (18.46), sont traditionnellement désignées par la lettre Y et portent le nom[17] d'*harmoniques sphériques*. Ainsi, les fonctions cherchées, $Y_{lm}(\theta, \phi)$, satisfont les deux équations aux dérivées partielles :

$$-\Big(\frac{\partial^2}{\partial\theta^2} + \cot\theta\,\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\Big)\,Y_{lm}(\theta,\,\phi) = l(l+1)\,Y_{lm}(\theta,\,\phi) \ , \tag{18.60}$$

[16] En relation avec les égalités (18.53), on note que l'opérateur (hermitique !) $\vec{L}^2$ est donné par une combinaison à coefficients *réels* d'opérateurs différentiels du *second* ordre.

[17] Cette appellation provient du fait que les Y_{lm} sont des fonctions harmoniques particulières (une fonction est dite *harmonique* si son Laplacien est nul). Pour cette raison, on peut aussi définir des harmoniques *circulaires* dans le plan $\mathbb{R}^2$, ou dans $\mathbb{R}^D$ pour n'importe quelle dimension $D > 1$. Dans le plan, les fonctions harmoniques jouent un rôle essentiel puisque, par les conditions de Cauchy - Riemann, les parties réelle et imaginaire d'une fonction holomorphe sont harmoniques.

$$-\mathrm{i}\frac{\partial}{\partial\phi}Y_{lm}(\theta,\,\phi)=m\,Y_{lm}(\theta,\,\phi)\ . \tag{18.61}$$

Dans la théorie générale du moment cinétique, notamment lors de la mise en œuvre de la procédure d'échelle, il a toujours été supposé que les vecteurs propres sont de norme *finie* ; cette propriété, jointe aux relations caractéristiques de commutation et au confinement des variables angulaires, est à l'origine de la *quantification* des nombres l et m ; dans la représentation en cours, ceci impose aux fonctions Y_{lm} d'être de module carré sommable. En conséquence, seules les solutions des équations précédentes possédant cette propriété – et d'autres d'ailleurs, assurant notamment qu'aucune moyenne d'observable n'est divergente – doivent être retenues physiquement ; normalisées conventionnellement à l'unité, elles doivent ainsi satisfaire[18] :

$$\boxed{\int_0^{\pi}\sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,|Y_{lm}(\theta,\,\phi)|^2=1} \tag{18.62}$$

Comme on l'a vu à de multiples reprises dans le Tome I, une telle condition effectue le tri dans l'ensemble des solutions des équations (18.60) et (18.61). Toutefois, elle n'est pas la seule à effectuer un passage au crible de ces solutions ; en effet, la fonction d'onde d'une particule (sans spin) doit avoir une et une seule valeur en un point donné de l'espace, elle doit être monovaluée[19] et continue ; ceci impose notamment qu'elle prenne la *même* valeur en ϕ et en $\phi+2\pi$, d'où une autre condition à satisfaire :

$$Y_{lm}(\theta,\,\phi)=Y_{lm}(\theta,\,\phi+2\pi)\ . \tag{18.63}$$

D'ailleurs, on sait du chapitre 17 que pour toute rotation $\mathcal{R}$ du système dans l'espace $\mathbb{R}^3$, la fonction d'onde $\Psi(\vec{r},\,t)$ se transforme en $\Psi(\mathcal{R}^{-1}\vec{r},\,t)$ (voir (17.93)), sans aucune phase additionnelle.

Ces conditions aux limites étant précisées, la résolution pragmatique des équations aux dérivées partielles peut être entreprise. La plus simple est l'équation (18.61), qui ne contient que l'opérateur différentiel $\frac{\partial}{\partial\phi}$; elle s'intègre immédiatement et a pour solution générale :

$$Y_{lm}(\theta,\,\phi)=\Theta_{lm}(\theta)\,\mathrm{e}^{\mathrm{i}m\phi}\ , \tag{18.64}$$

où Θ_{lm} est pour l'instant une fonction quelconque de θ jouant le rôle d'une constante d'intégration vis-à-vis de ϕ. Proportionnelle à $\mathrm{e}^{\mathrm{i}m\phi}$, la fonction Y_{lm} n'est *monovaluée* que si m est un *entier* ($m=0,\pm1,\pm2,\ldots$). De la théorie générale, on savait que m pouvait être entier *ou* demi-entier et que l et m sont solidaires dans cette double possibilité ; en fait, compte tenu du lien entre les rotations spatiales et le moment cinétique orbital, m

[18] Certains auteurs utilisent aussi la "mesure" $\mathrm{d}\mu\stackrel{\text{déf}}{=}[1/(4\pi)]\sin\theta\mathrm{d}\theta\mathrm{d}\phi$, qui représente l'angle solide "normalisé" $\mathrm{d}\tilde{\Omega}\stackrel{\text{déf}}{=}\mathrm{d}\Omega/(4\pi)$. La normalisation s'écrit alors $\int\mathrm{d}\mu\,|\tilde{Y}_{lm}|^2=1$. Les $\tilde{Y}_{lm}$ et les Y_{lm} diffèrent d'un facteur $\sqrt{4\pi}$. Avec cette mesure, $\langle 1\rangle=1$, $1/(4\pi)$ étant finalement la densité de probabilité uniforme d'un angle solide (dans $\mathbb{R}^2$, la même mesure uniforme est $1/(2\pi)$: à chaque fois, le dénominateur est la surface Ω_D de la "'sphère" de rayon 1 dans $\mathbb{R}^D$ – voir la note 87 du chapitre 19).

[19] Si on s'autorisait une phase entre $\Phi(0)$ et $\Phi(2\pi)$, les prévisions physiques resteraient encore inaltérées, mais la fonction d'onde ne serait pas continue. Sur le caractère monovalué de la fonction d'onde, voir [4] p. 179 et [29]. Au chapitre 20, on verra que la fonction d'onde (à deux composantes) d'un spin 1/2 acquiert une phase $\mathrm{e}^{\mathrm{i}\pi}=-1$ lors d'une rotation de 2π : avec $S=1/2$, $\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}(2\pi)(\hbar/2)}=-1$.

ne peut être ici qu'un entier, et il en va alors de même pour le nombre quantique l. On retrouve ici, une fois encore (voir Tome I), le rôle décisif des conditions que doit satisfaire la fonction d'onde sur l'apparition précise de la quantification. Par ailleurs, la possibilité unique pour l d'être un entier (et non un demi-entier) est clairement le résultat d'avoir pu écrire $\vec{L}$ en terme des deux grandeurs fondamentales $\vec{r}$ et $\vec{p}$, et d'avoir représenté ces dernières suivant les prescriptions de Schrödinger. Rien de tel n'existe pour un moment cinétique de spin, pour lequel les valeurs entières *ou* demi-entières sont possibles.

Ces résultats étant acquis, on peut poursuivre la résolution de l'équation aux valeurs propres (18.60) soit en se branchant sur la littérature de Mathématiques appliquées (équation de Legendre[20], voir p. 735 quelques éléments techniques à ce sujet), soit en utilisant directement les résultats établis ci-dessus, se contentant d'une résolution *a minima* très suffisante pour la plupart des applications. C'est ce qui va être fait maintenant, en réitérant avec les fonctions Y_{lm} la procédure d'échelle développée antérieurement sur les *kets* $|j\,m\rangle$, celles-ci en constituant la représentation-q avec le glissement de notation $j \rightarrow l$: $Y_{lm}(\theta,\,\phi) \equiv \langle\theta,\,\phi|Y_{lm}\rangle$.

La valeur maximale m_{max} de m étant égale à l, de la théorie générale pour $\vec{J}$, et d'après (18.23), L_+ agissant sur $Y_{lm_{\text{max}}=l}$ doit donner zéro :

$$L_+\,Y_{ll}(\theta,\,\phi) = 0\ . \tag{18.65}$$

Compte tenu de l'expression (18.59) de l'opérateur L_+, cette dernière équation s'écrit :

$$L_+\,Y_{ll}(\theta,\,\phi) \equiv \hbar\,\mathrm{e}^{+\mathrm{i}\phi}\Big(\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\,\frac{\partial}{\partial\phi}\Big)\,\Theta_{ll}(\theta)\,\mathrm{e}^{\mathrm{i}l\phi} = 0\ . \tag{18.66}$$

L'opérateur ∂_ϕ agissant sur $\mathrm{e}^{\mathrm{i}l\phi}$ donne un simple facteur $\mathrm{i}l$; après simplification par $\mathrm{e}^{\mathrm{i}l\phi}$, il reste une équation où figure la seule variable θ :

$$\Big(\frac{\mathrm{d}}{\mathrm{d}\theta} - l\,\cot\theta\Big)\,\Theta_{ll}(\theta) = 0 \quad\Longleftrightarrow\quad \frac{1}{\Theta_{ll}}\,\frac{\mathrm{d}\Theta_{ll}}{\mathrm{d}\theta} = l\,\frac{\mathrm{d}}{\mathrm{d}\theta}\,\ln(\sin\theta)\ ; \tag{18.67}$$

la deuxième écriture venant du fait que $\cot\theta$ est la dérivée[21] de $\ln(\sin\theta)$. L'intégration est immédiate et fournit $\Theta_{ll}(\theta) = C_l\,(\sin\theta)^l$, où C_l est une constante arbitraire, qui sera fixée par normalisation. Au total :

$$\boxed{Y_{ll}(\theta,\phi) = C_l\,(\sin\theta)^l\,\mathrm{e}^{\mathrm{i}l\phi}} \tag{18.68}$$

Ayant obtenu la représentation-q de $|Y_{lm_{\text{max}}=l}\rangle$, il est maintenant possible de construire les $2l$ autres fonctions Y_{lm}, avec $m = l-1, l-2, \ldots, -l$, par application répétée de L_- et usage de la relation générale (18.38) qui, lue à l'envers s'écrit ici :

$$|Y_{lm-1}\rangle = \frac{1}{\hbar\,\sqrt{l(l+1) - m(m-1)}}\,L_-|Y_{lm}\rangle\ . \tag{18.69}$$

[20] Outre l'équation différentielle qui porte son nom, et rencontrée dans ses recherches en Mécanique, Adrien-Marie Legendre (1752-1833) est l'auteur de travaux importants en Théorie des nombres (démonstration de l'irrationnalité de π) et en Analyse (fonctions elliptiques). On lui doit aussi la méthode des moindres carrés (1805), juste avant Gauss (1809). Il introduisit le symbole ∂ dans son mémoire intitulé *"Mémoire sur la manière de distinguer les maxima des minima dans le Calcul des Variations"* (1786).

[21] $\cot\theta$ est plus précisément la dérivée $\ln|\sin\theta|$, mais θ étant compris entre 0 et π, le module est inutile.

En représentation-q, compte tenu de (18.59) et de $Y_{lm}(\theta,\,\phi) = \Theta_{lm}(\theta)\mathrm{e}^{\mathrm{i}m\phi}$, cette relation prend la forme :

$$Y_{lm-1}(\theta,\,\phi) = \frac{\mathrm{e}^{-\mathrm{i}\phi}}{\sqrt{l(l+1)-m(m-1)}}\Big(-\frac{\partial}{\partial\theta} - m\cot\theta\Big)\,Y_{lm}(\theta,\,\phi)\ . \qquad (18.70)$$

Ainsi par exemple, pour $m = l$ et selon (18.68) :

$$Y_{ll-1}(\theta,\,\phi) = \frac{\mathrm{e}^{-\mathrm{i}\phi}}{\sqrt{2l}}\Big(-\frac{\partial}{\partial\theta} - l\cot\theta\Big)Y_{ll}(\theta,\,\phi) = -C_l\sqrt{2l}\,\mathrm{e}^{\mathrm{i}(l-1)\phi}\cos\theta(\sin\theta)^{l-1}\ ; \qquad (18.71)$$

on note que la dépendance en ϕ de $Y_{lm=l-1}$ est bien $\mathrm{e}^{\mathrm{i}m\phi}$. Connaissant maintenant Y_{ll-1}, on peut construire Y_{ll-2} de la même façon, et ainsi de suite. Il n'est pas difficile d'obtenir des formules générales[22] pour Y_{ll-k} ; il apparaît de fait des fonctions connues appelées[23] *polynômes associés de Legendre*, dont la définition et certaines propriétés seront données en temps utile (voir p. 735). Compte tenu de la réalité de l'opérateur différentiel agissant sur θ, et du fait que le facteur $\mathrm{e}^{\mathrm{i}m\phi}$ apparaît nécessairement, on voit que les harmoniques sphériques sont de la forme :

$$Y_{lm}(\theta,\,\phi) = \tilde{C}_{lm}\mathrm{e}^{\mathrm{i}m\phi}\mathcal{P}_{lm}(\theta) \qquad (18.72)$$

où $\mathcal{P}_{lm}(\theta)$ est un polynôme réel en $\sin\theta$ et $\cos\theta$; toute autre phase conventionnelle liée à la normalisation est donc exclusivement dans la constante $\tilde{C}_{lm}$.

Les fonctions Y_{lm} sont orthogonales entre elles puisqu'elles sont fonctions propres du couple d'opérateurs hermitiques $(\vec{L}^2,\,L_z)$, associées à des couples distincts de valeurs propres ; après normalisation à l'unité de chaque Y_{lm}, on a donc[24] :

$$\langle Y_{lm}|Y_{l'm'}\rangle \equiv \int_0^{\pi}\sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\;Y_{lm}^*(\theta,\,\phi)\,Y_{l'm'}(\theta,\,\phi) = \delta_{ll'}\,\delta_{mm'}\ . \qquad (18.73)$$

Comme d'habitude, une telle contrainte laisse encore un facteur de phase indéterminé, sans aucune incidence physique. La convention d'usage consiste à prendre $Y_{l0}(0,0)$ réel et positif ; alors $Y_{lm=0}(0,\,0) = [(2l+1)/(4\pi)]^{1/2}$ et $Y_{lm\neq 0}(0,\,0) = 0$, soit (voir note 18) :

$$Y_{lm}(0,\,0) = \delta_{m0}\sqrt{\frac{2l+1}{4\pi}} \equiv \sqrt{2l+1}\;\tilde{Y}_{lm}(0,\,0)\ . \qquad (18.74)$$

La constante $\tilde{C}_{l0}$ étant conventionellement prise réelle, il s'ensuit que la fonction $Y_{l0}(\theta,\,\phi)$ l'est aussi ; les phases de toutes les autres Y_{lm} de *même* l s'en trouvent automatiquement fixées grâce aux relations (18.69).

En outre, les Y_{lm} constituent une base complète orthonormée pour les fonctions de carré sommable définies de la sphère unité de $\mathbb{R}^3$ dans $\mathbb{C}$. À ce titre, elles satisfont

[22] voir [15], p. 446, [25], complément $\mathrm{A_{VI}}$.

[23] La terminologie est source de confusion : $\mathcal{P}_{lm}(\theta)$ est bien un polynôme à deux variables ($\cos\theta$ et $\sin\theta$, mais $P_l^{|m|}(\cos\theta) \stackrel{\text{déf}}{=} \mathcal{P}_{lm}(\theta)$ n'en est un que si m est pair (voir (18.117)).

[24] bien noter que l'élément différentiel n'est pas simplement $\mathrm{d}\theta\,\mathrm{d}\phi$.

la relation de fermeture exprimant cette complétude ; comme l'intégration sur θ se fait avec $\sin\theta\,\mathrm{d}\theta = -\mathrm{d}(\cos\theta)$, la relation de fermeture s'écrit explicitement[25] :

$$\sum_{l=0}^{+\infty}\sum_{m=-l}^{+l} Y_{lm}(\theta,\phi)\,Y_{lm}^*(\theta',\phi') = \delta(\cos\theta-\cos\theta')\,\delta(\phi-\phi')\ , \tag{18.76}$$

soit :

$$\boxed{\sum_{l=0}^{+\infty}\sum_{m=-l}^{+l} Y_{lm}(\theta,\phi)\,Y_{lm}^*(\theta',\phi') = \frac{1}{\sin\theta}\delta(\theta-\theta')\,\delta(\phi-\phi')} \tag{18.77}$$

où il serait à nouveau superflu d'écrire $|\sin\theta|$ puisque $\theta\in[0,\,\pi]$. Ainsi toute fonction de carré sommable $f(\theta,\phi)$ peut être développée en série des Y_{lm} :

$$\boxed{f(\theta,\phi) = \sum_{l=0}^{+\infty}\sum_{m=-l}^{+l} f_{lm}\,Y_{lm}(\theta,\phi)} \tag{18.78}$$

les coefficients f_{lm} s'expriment comme d'habitude par un produit scalaire, grâce à l'orthonormalisation (18.73) des Y_{lm} ; l'intégrale représentant toujours ce produit scalaire, on a :

$$\boxed{f_{lm} = \langle Y_{lm}|f\rangle = \int_0^{\pi}\sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\ Y_{lm}^*(\theta,\phi)\,f(\theta,\phi)} \tag{18.79}$$

La parité (ou inversion d'espace), Π, change θ en $\pi-\theta$ et ϕ en $\phi+\pi$, puisqu'elle est égale au produit (dans un ordre quelconque) d'une symétrie-miroir par rapport à $x\mathrm{O}y$ et d'une rotation de 180° autour de $\mathrm{O}z$; par cette transformation, les Y_{lm} au plus changent de signe : ce sont des fonctions de parité déterminée. En effet, puisque l'opérateur parité Π commute avec $\vec{L}$ et que le couple (l, m) spécifie complètement l'état propre de $(\vec{L}^2, L_z)$ (il n'y a pas de dégénérescence), le vecteur $|Y_{lm}\rangle$ est forcément propre[26] de Π ; comme les valeurs propres de Π sont $\varpi=\pm1$, on a $\Pi|Y_{lm}\rangle = \pm|Y_{lm}\rangle$. Par ailleurs $[\Pi, L_\pm]=0$, donc $|Y_{lm}\rangle$ et $|Y_{lm\pm1}\rangle$ ont la même valeur propre ϖ : tous les états $\{|Y_{lm}\rangle\}_{-l\leq m\leq +l}$ ont donc la même parité, qui est celle de $Y_{ll} = C_l(\sin\theta)^l\,\mathrm{e}^{\mathrm{i}l\phi}$. Dans l'inversion d'espace, $\sin\theta$ est invariant et il vient seulement une phase $\mathrm{e}^{\mathrm{i}l\pi}=(-1)^l$. Finalement :

$$\boxed{\Pi\,Y_{lm}(\theta,\phi) \stackrel{\text{déf}}{=} Y_{lm}(\pi-\theta,\phi+\pi) = (-1)^l Y_{lm}(\theta,\phi)} \tag{18.80}$$

À titre d'exemple des résultats précédents, donnons une relation utile dans les applications. Soit la fonction $\cos\theta\,Y_{lm}(\theta,\phi)$, qui a la parité $(-1)^{l+1}$; comme les $Y_{lm}(\theta,\phi)$

[25] On utilise :

$$\delta(f(x)) = \frac{1}{|f'(x_0)|}\,\delta(x-x_0)\ , \tag{18.75}$$

relation qui est vraie quand la fonction s'annule une seule fois, en x_0 – et se généralise immédiatement quand f a plusieurs zéros. Ici, comme θ et θ' sont compris entre 0 et π, à θ fixé, l'équation $\cos\theta'-\cos\theta=0$ n'a qu'une seule solution $\theta'=\theta$.

[26] Cet argument a déjà été plusieurs fois utilisé, notamment pour prévoir que les fonctions propres de l'oscillateur harmonique (linéaire) sont forcément paires *ou* impaires (voir aussi section 17.6).

forment une base compète, on peut *a priori* écrire :

$$\cos\theta\, Y_{lm}(\theta,\phi) = \sum_{l',m'} A_{l'm'} Y_{l'm'}(\theta,\phi) \ ; \tag{18.81}$$

en fait, seule la valeur $m' = m$ figure à droite puisque la dépendance en ϕ est de la forme $\mathrm{e}^{\mathrm{i}m\phi}$. Par ailleurs, la parité impose que seuls les l' de parité opposée à l donnent un coefficient $A_{l'm}$ non nul. Il suffit donc de calculer les intégrales :

$$A_{l'm} = \langle Y_{l'm} | \cos\theta\, Y_{lm}\rangle = \int_0^{\pi} Y^*_{l'm'}(\theta,\phi)\cos\theta\, Y_{lm}(\theta,\phi)\,\sin\theta \mathrm{d}\theta\mathrm{d}\phi \ ; \tag{18.82}$$

on trouve que ces intégrales sont nulles sauf si $l' = \pm 1$, de sorte que finalement :

$$\boxed{\cos\theta\, Y_{lm}(\theta,\phi) = A_{l+1m}Y_{l+1m}(\theta,\phi) + A_{l-1m}Y_{l-1m}(\theta,\phi)} \tag{18.83}$$

où les coefficients $A_{l\pm 1m}$ ont pour expression :

$$A_{l\pm 1m} = \sqrt{\frac{(l+\frac{1}{2}\pm\frac{1}{2}+m)(l+\frac{1}{2}\pm\frac{1}{2}-m)}{(2l\pm 1)(2l+2\pm 1)}} \ . \tag{18.84}$$

Il n'est pas non plus difficile d'établir le mode de transformation de la fonction Y_{lm} par la conjugaison complexe, qui doit être proportionnelle[27] à $Y_{l\,-m}$; cela étant précisé, introduisons une phase α_m telle que :

$$[Y_{lm}(\theta,\phi)]^* = \mathrm{e}^{\mathrm{i}\alpha_m} Y_{l\,-m}(\theta,\phi) \ . \tag{18.85}$$

Revenons maintenant à (18.38), récrite comme :

$$L_- Y_{lm}(\theta,\,\phi) = \hbar\sqrt{l(l+1)-m(m-1)}\, Y_{lm-1}(\theta,\,\phi) \tag{18.86}$$

et appliquons-lui l'opération K de renversement du temps, qui inverse les moments cinétiques et conjugue les fonctions :

$$-L_- Y^*_{lm}(\theta,\,\phi) = \hbar\sqrt{l(l+1)-m(m-1)}\, Y^*_{lm-1}(\theta,\,\phi) \ ; \tag{18.87}$$

compte tenu de la définition de la phase α_m dans (18.85), on en déduit $\mathrm{e}^{\mathrm{i}\alpha_{m-1}} = -\mathrm{e}^{\mathrm{i}\alpha_m}$, soit $\alpha_m = \alpha_{m-1} + \pi$. Comme $\alpha_0 = 0$ puisque $Y_{l0}(\theta,\,\phi)$ est réel, il vient $\alpha_m = m\pi$, soit :

$$\boxed{KY_{lm}(\theta,\phi) = Y^*_{lm}(\theta,\phi) = (-1)^m Y_{l\,-m}(\theta,\phi)} \tag{18.88}$$

On retrouve à ce propos la particularité annoncée au chapitre 17 p. 700 : tant qu'il n'est question que de moments cinétiques *orbitaux* (par opposition à des spins), le carré de K est l'identité puisque $(-1)^{2m} = 1$ quel que soit m entier.

[27] La conjugaison complexe change m en $-m$, en accord avec le renversement du temps, qui inverse les moments cinétiques (voir (17.186)).

▼ À propos de l'équation de Legendre

Dans un double but, donnons quelques compléments techniques à propos des solutions acceptables de l'équation (18.60). D'une part, et pour usage ultérieur, afin de préciser la relation entre les harmoniques sphériques et des fonctions spéciales répertoriées, appelées fonctions associées de Legendre P_l^m ([30], p. 512-513) qui, pour l et m entiers, se réduisent en fait à des *polynômes* ; d'autre part pour retrouver le scénario maintenant classique où les conditions aux limites jouent un rôle majeur pour l'apparition de la quantification.

L'équation aux dérivées partielles (18.60) se résout par séparation des variables ; en effet, en posant :

$$Y_{lm}(\theta, \phi) = \Phi(\phi)\,\Theta(\theta) \ , \tag{18.89}$$

et en reportant, il vient :

$$-\Phi(\phi)\Big(\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} + \cot\theta\frac{\mathrm{d}}{\mathrm{d}\theta}\Big)\Theta(\theta) - \Theta(\theta)\frac{1}{\sin^2\theta}\frac{\mathrm{d}^2}{\mathrm{d}\phi^2}\Phi(\phi) = l(l+1)\Phi(\phi)\Theta(\theta) \ , \tag{18.90}$$

soit :

$$-\frac{1}{\Phi}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\phi^2} = \frac{\sin^2\theta}{\Theta}\Big(\frac{\mathrm{d}^2\Theta}{\mathrm{d}\theta^2} + \cot\theta\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\Big) + l(l+1)\sin^2\theta \ . \tag{18.91}$$

Cette dernière équation montre clairement la séparation des variables : le premier membre ne dépend que de ϕ, le second terme ne dépend que de θ. Chacun d'entre eux doit donc être une seule et même constante, notée C :

$$-\frac{1}{\Phi}\frac{\mathrm{d}^2\Phi}{\mathrm{d}\phi^2} = C = \frac{\sin^2\theta}{\Theta}\Big(\frac{\mathrm{d}^2\Theta}{\mathrm{d}\theta^2} + \cot\theta\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\Big) + l(l+1)\sin^2\theta \ . \tag{18.92}$$

Par ailleurs, en dérivant une fois de plus (18.61), on obtient :

$$\frac{\partial^2}{\partial\phi^2}Y_{lm}(\theta,\phi) = -m^2Y_{lm}(\theta,\phi) \quad\Longleftrightarrow\quad \frac{\mathrm{d}^2}{\mathrm{d}\phi^2}\Phi(\phi) = -m^2\Phi(\phi) \ . \tag{18.93}$$

La constante C étant ainsi égale à m^2, carré du nombre quantique magnétique m, l'équation en Θ est :

$$\Big(\frac{\mathrm{d}^2\Theta}{\mathrm{d}\theta^2} + \cot\theta\,\frac{\mathrm{d}\Theta}{\mathrm{d}\theta}\Big) - \frac{m^2}{\sin^2\theta}\,\Theta + l(l+1)\,\Theta = 0 \ . \tag{18.94}$$

Le changement de variable $\cos\theta = X$, $\Theta(\theta) \stackrel{\text{déf}}{=} F(X)$ transforme (18.94) en :

$$(1-X^2)\,\frac{\mathrm{d}^2F}{\mathrm{d}X^2} - 2X\,\frac{\mathrm{d}F}{\mathrm{d}X} + \Big[l(l+1) - \frac{m^2}{1-X^2}\Big]\,F = 0 \ ; \tag{18.95}$$

avec $m = 0$, (18.95) est une équation dûment répertoriée, appelée *équation de Legendre*, que l'on peut récrire sous la forme de Sturm - Liouville[28] :

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}X}\Big[(1-X^2)\,\frac{\mathrm{d}F}{\mathrm{d}X}\Big] + l(l+1)\,F = 0} \tag{18.97}$$

[28]On appelle équation de Sturm - Liouville une équation différentielle du second ordre de la forme :

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big[a_2(x)\frac{\mathrm{d}f}{\mathrm{d}x}\Big] + a_0(x)f(x) = 0 \ ; \tag{18.96}$$

avec $a_2(x) = \mathrm{C}^{\mathrm{ste}}$ et $a_0(x) = \varepsilon - U(x)$, on obtient l'équation aux valeurs propres pour une particule sur $\mathbb{R}$ (voir Tome I, éq. (15.5)).

L'équation (18.95) est l'*équation associée de Legendre.*

Le mouvement lié à une dimension n'est jamais dégénéré (voir Tome I, section (15.1)) ; l'équation (18.97) étant invariante si X est changé en $-X$, c'est-à-dire si θ est changé en $\pi-\theta$, on peut se borner à trouver des solutions qui sont soit paires soit impaires dans ce changement (correspondant à une symétrie-miroir par rapport au plan xOy).

L'équation de Legendre (18.97) possède deux solutions linéairement indépendantes permettant d'engendrer la solution la plus générale par combinaison linéaire. Toutefois, quel que soit l, l'une de ces deux solutions n'est pas acceptable physiquement car elle présente des singularités logarithmiques en $X = \pm 1$. Afin de mettre celles-ci en évidence le plus simplement possible, considérons le voisinage de $X = 1$; l'équation (18.95) avec $m = 0$ a alors la forme approximative :

$$2(1-X)\,\frac{\mathrm{d}^2F}{\mathrm{d}X^2} - 2\,\frac{\mathrm{d}F}{\mathrm{d}X} + l(l+1)\,F \simeq 0 \ . \tag{18.98}$$

Reportant une fonction $f(X)$ qui se comporte comme $\ln(1-X)$ dans ce voisinage :

$$f(X) = \ln(1-X) \ , \qquad f'(X) = -\frac{1}{1-X} \ , \qquad f''(X) = -\frac{1}{(1-X)^2} \ , \tag{18.99}$$

il vient :

$$-2\,\frac{1}{1-X} + 2\,\frac{1}{1-X} + l(l+1)\,\mathcal{O}[\ln(1-X)] \sim 0 \ . \tag{18.100}$$

Les termes les plus divergents se compensent bien, le logarithme dans (18.100) étant un infiniment grand d'ordre inférieur. Ainsi, l'équation de Legendre a des solutions qui, pour $X \simeq 1$, se comportent à peu près comme $F(X) \simeq \ln(1-X)$, $\Theta(\theta) \simeq \ln(1-\cos\theta)$.

Montrons que de telles solutions doivent être rejetées sur des considérations physiques. La première idée venant à l'esprit (presque par habitude) est d'exploiter la condition de normalisabilité sur Θ, qui introduit forcément quelque part une contribution intégrale du genre $\int_{X_0}^{+1}\mathrm{d}X\,[\ln(1-X)]^2$; en fait, comme cette intégrale existe[29], ce n'est pas la condition à invoquer pour faire le tri. Une bonne raison est en fait la suivante : l'opérateur donnant l'énergie cinétique de rotation contient des dérivées secondes en θ, qui introduit des intégrales du genre $\int_{-1}^{+1}\mathrm{d}X\,[\ln(1-X)]/(1-X)^2$ qui sont, elles, visiblement divergentes ; ces solutions à singularités logarithmiques ne sont donc pas acceptables.

Les solutions régulières, elles, peuvent être cherchées sous la forme d'une série entière :

$$F(X) = \sum_{k=0}^{+\infty} c_k\,X^k \ , \tag{18.101}$$

et la substitution dans (18.97) produit la relation de récurrence :

$$(k+1)(k+2)\,c_{k+2} = [k(k+1) - l(l+1)]\,c_k \ , \tag{18.102}$$

[29] L'intégrale $\int_0^a \mathrm{d}x\,(\ln x)^2$, $a > 0$, existe ; on peut s'en convaincre en effectuant des intégrations par parties. On sait d'ailleurs l'exprimer à l'aide de fonctions élémentaires : elle vaut $a[(\ln a - 1)^2 + 1]$.

qui relie les coefficients de deux en deux, et nécessite donc de se donner les deux premiers d'entre eux : si on prend $c_0 \neq 0$ et $c_1 = 0$, on obtient une fonction paire ; à l'inverse, le choix $c_1 \neq 0$ et $c_0 = 0$ produit une fonction impaire. Dans les deux cas, la fonction est déterminée comme d'habitude à un coefficient multiplicatif global près.

Oublions un instant que l est un entier. Le rapport c_{k+2}/c_k est $\simeq k/(k+2)$ quand $k \gg 1$: il en résulte que, en $X = 1$ (soit pour $\theta = 0$), la série (18.101) se comporte essentiellement comme la série harmonique, qui est divergente. Pour éliminer cette singularité, il faut impérativement que la série s'arrête à un certain rang k, une condition nécessaire et suffisante étant que l soit un entier naturel : alors, par la récurrence (18.102), tous les coefficients $c_k + 2$ sont nuls si $k \geq l$, et la série se réduit à un polynôme pour lequel aucune question de convergence ne se pose. On retrouve ainsi une fois de plus la nécessaire quantification, ici du moment cinétique, en conséquence des conditions imposées par le sens physique attribué à toute fonction d'onde.

Précisons maintenant les relations entre les harmoniques sphériques et les bonnes solutions de l'équation associée de Legendre (18.95) qui, puisque l est entier, sont des polynômes, traditionnellement notés P_l^m et appelés *polynômes associés de Legendre*. Pour $m \geq 0$, ceux-ci s'obtiennent à partir des *polynômes de Legendre* $P_l(X)$, solutions de (18.97) avec l entier. En fixant conventionnellement[30] la valeur de ces derniers en $X = 1$, soit $P_l(1) = 1$, ils s'obtiennent à partir de la formule de Rodrigues[31] :

$$\boxed{P_l(X) \overset{\text{déf}}{=} \frac{(-1)^l}{2^l\, l!}\, \frac{\mathrm{d}^l}{\mathrm{d}X^l}\, (1 - X^2)^l} \tag{18.103}$$

sur laquelle il est visible que $P_l(X)$ est un polynôme de degré l en X, pair si l est pair et inversement : $P_l(-X) = (-1)^l P_l(X)$. Par la formule de Cauchy, la formule de Rodrigues peut se transcrire en terme d'une intégrale complexe (intégrale de Schäfli) :

$$P_l(z) = \frac{(-1)^l}{2^l}\, \frac{1}{2\mathrm{i}\pi} \int_\Gamma \frac{(1-\xi^2)^l}{(\xi - z)^{l+1}}\, \mathrm{d}\xi \tag{18.104}$$

où Γ est une boucle entourant le point $z \in \mathbb{C}$, parcourue une fois et une seule dans le sens positif.

On peut aussi construire les $P_l(X)$ de proche en proche grâce à la relation de récurrence :

$$(l+1)P_{l+1}(X) = (2l+1)XP_l(X) - lP_{l-1}(X) \ . \tag{18.105}$$

Ces polynômes satisfont également une relation utile à connaître pour les intégrations :

$$\frac{\mathrm{d}}{\mathrm{d}X}[P_{l+1}(X) - P_{l-1}(X)] = (2l+1)P_l(X) \ . \tag{18.106}$$

Enfin, on peut montrer que tous leurs zéros sont réels, et sont dans l'intervalle $[-1, +1]$.

[30] L'équation différentielle de Legendre étant homogène, ses solutions sont définies à un facteur près.

[31] Benjamin-Olinde Rodrigues-Henriques (1795-1851) a publié cette formule en 1816 dans un article intitulé *"Mémoire sur l'attraction des sphéroïdes"*, Correspondance de l'École Royale Polytechnique, **3**, 361 (1816) (voir [31]). Pour la démonstration de la formule de Rodrigues, voir problème 18.7.13.

Une fonction génératrice de ces polynômes est[32] :

$$F(X,t) \stackrel{\text{déf}}{=} \sum_{l=0}^{+\infty} t^l\, P_l(X) \quad (|t|<1) \ , \qquad P_l(X) = \frac{1}{l!}\Big(\frac{\partial^l F}{\partial t^l}\Big)_{t=0} \tag{18.109}$$

et a pour expression :

$$F(X,t) = \frac{1}{\sqrt{1-2Xt+t^2}} \ . \tag{18.110}$$

Pour tout X réel fini, la fonction $(1-2Xz+z^2)^{1/2}$ a deux points de branchement en $z_\pm = X \pm (X^2-1)^{1/2}$, où la fonction $(X^2-1)^{1/2}$ peut être définie comme la détermination prenant des valeurs réelles positives sur $|X|>1$. On peut toujours choisir les coupures de $(1-2Xz+z^2)^{1/2}$ de sorte qu'il existe un voisinage de l'origine où cette fonction est holomorphe ; par le théorème de Laurent (ou la formule de Cauchy), on peut ainsi écrire une autre représentation des polynômes P_l, souvent utile dans les applications :

$$P_l(X) = \frac{1}{2\mathrm{i}\pi}\oint_\Gamma \frac{1}{z^{l+1}}\frac{1}{(1-2Xz+z^2)^{1/2}}\,\mathrm{d}z \ , \tag{18.111}$$

où Γ est une boucle entourant l'origine de $\mathbb{C}$ une fois et une seule, parcourue dans le sens trigonométrique positif[33].

Les polynômes P_l sont orthogonaux suivant (voir problème 18.7.13) :

$$\boxed{\int_{-1}^{+1} P_l(X)\,P_{l'}(X)\,\mathrm{d}X \equiv \int_0^\pi P_l(\cos\theta)\,P_{l'}(\cos\theta)\,\sin\theta\,\mathrm{d}\theta = \frac{2}{2l+1}\,\delta_{ll'}} \tag{18.112}$$

En outre, ils apparaissent dans un développement souvent utilisé en Physique. Soit deux points de $\mathbb{R}^3$, extrémités de deux vecteurs $\vec{r}_1$ et $\vec{r}_2$, et notons r_{12} la distance entre ces deux points : $r_{12} \stackrel{\text{déf}}{=} \|\vec{r}_1 - \vec{r}_2\|$. On a alors :

$$\frac{1}{r_{12}} = \frac{1}{\sqrt{r_1^2 - 2r_1r_2\cos\gamma + r_2^2}} = \frac{1}{r_>}\frac{1}{\sqrt{1-2\frac{r_<}{r_>}\cos\gamma + \left(\frac{r_<}{r_>}\right)^2}} \tag{18.113}$$

[32]La notion de fonction génératrice a été rencontrée antérieurement à propos de l'oscillateur harmonique (voir Tome I, éq. (16.48)). La définition adoptée ici se conforme aux habitudes en la matière, et n'exhibe pas le facteur $\frac{1}{l!}$ dans la série (18.109), incorporé d'emblée dans la définition usuelle (18.103) des $P_l(X)$. Plus généralement, la fonction $(1-2Xt+t^2)^{-\alpha}$ – où l'exposant positif α peut prendre diverses valeurs – définit plusieurs classes de polynômes remarquables.

On connaît une autre fonction génératrice :

$$\mathrm{e}^{Xt}\,J_0(t\sqrt{1-X^2}) = \sum_{l=0}^{+\infty} P_l(X)\frac{t^l}{l!} \ , \tag{18.107}$$

J_0 étant la fonction de Bessel ordinaire d'indice nul, d'où une autre expression des P_l :

$$P_l(X) = \frac{\partial^l}{\partial t^l}\left[\mathrm{e}^{Xt}\,J_0(t\sqrt{1-X^2})\right]_{t=0} \ . \tag{18.108}$$

[33]Avec $X=\cos\theta$, les deux points de branchement de $(1-2Xz+z^2)^{1/2}$ sont $z_\pm = \mathrm{e}^{\pm\mathrm{i}\theta}$; dans ces conditions, la boucle Γ, quoique de forme libre, doit toutefois laisser ces deux points en son extérieur.

où $r_< = \inf(r_1, r_2)$, $r_> = \sup(r_1, r_2)$ et $\gamma = (\vec{r}_1, \vec{r}_2)$. Par comparaison avec (18.109) et (18.110) avec $X = \cos\gamma$ et $t = \frac{r_<}{r_>}$, on obtient :

$$\boxed{\frac{1}{r_{12}} = \sum_{l\in\mathbb{N}} \frac{r_<^l}{r_>^{l+1}}\, P_l(\cos\gamma)} \tag{18.114}$$

Cette formule est très utile, par exemple, pour calculer l'énergie d'interaction de deux charges nues dans le vide, ou la valeur moyenne de l'énergie d'interaction de Coulomb pour deux électrons d'un atome (voir notamment chapitre 27, section 27.7).

Les premiers polynômes de Legendre sont :

$$P_0(X) = 1 \ , \ P_1(X) = X \ , \ P_2(X) = \frac{1}{2}(3X^2 - 1) \ , \ P_3(X) = \frac{1}{2}X(5X^2 - 3) \tag{18.115}$$

$$P_4(X) = \frac{1}{8}(35X^4 - 30X^2 + 3) \ , \quad P_5(X) = \frac{1}{8}X(63X^4 - 70X^2 + 15) \ . \tag{18.116}$$

Cela étant, les polynômes associés de Legendre $P_l^{m\geq 0}$, solutions analytiques de (18.95) en $X = 1$, peuvent s'obtenir comme suit[34] :

$$P_l^m(X) = (1 - X^2)^{m/2}\, \frac{\mathrm{d}^m}{\mathrm{d}X^m}\, P_l(X) \qquad (m \geq 0) \ ; \tag{18.117}$$

visiblement, $P_l^m(1) = 0$ si $m \neq 0$, de sorte que $P_l^m(1) = \delta_{m0} P_l(0) = \delta_{m0}$; $P_l^m(X)$ s'annule $l - m$ fois dans $]-1, +1[$ et deux fois en ± 1. Les polynômes associés de *même indice supérieur* m sont orthogonaux et satisfont ([4], éq. (9.62)) :

$$\int_{-1}^{+1} P_l^m(X)\, P_{l'}^m(X)\, \mathrm{d}X = \frac{2}{2l+1}\, \frac{(l+m)!}{(l-m)!}\, \delta_{ll'} \ . \tag{18.118}$$

Si l'indice supérieur n'est pas le même, les deux polynômes associés ne sont pas orthogonaux ; ils n'ont d'ailleurs aucune raison de l'être puisqu'ils sont solutions propres de deux Hamiltoniens *différents* (par l'entier m, voir (18.95)). Finalement, les harmoniques sphériques s'obtiennent comme :

$$\boxed{Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}\, \frac{(l-m)!}{(l+m)!}}\, (-1)^m\, \mathrm{e}^{\mathrm{i}m\phi}\, P_l^m(\cos\theta) \qquad (m \geq 0)} \tag{18.119}$$

Pour les valeurs négatives de m, on utilise $Y_{l\,-m} = (-1)^m\, Y_{lm}^*$ $(m > 0)$, d'où :

$$\boxed{Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}\, \frac{(l-|m|)!}{(l+|m|)!}}\, \mathrm{e}^{\mathrm{i}m\phi}\, P_l^{|m|}(\cos\theta) \qquad (m \leq 0)} \tag{18.120}$$

[34]Noter que $P_l^{m=0}(X) = P_l(X)$, et que seuls les P_l^m, $m \geq 0$ sont définis. Par ailleurs, comme déjà signalé, $P_l^m(X)$ n'est un "vrai" polynôme que si m est pair, comme il est visible sur l'égalité (18.117).

Enfin, les harmoniques sphériques satisfont une importante relation, dite *théorème d'addition* :

$$P_l(\cos\gamma) = \frac{4\pi}{2l+1}\sum_{m=-l}^{+l} Y_{lm}(\theta', \phi')\, Y_{lm}^*(\theta'', \phi'') \tag{18.121}$$

où (θ', ϕ') et (θ'', ϕ'') spécifient deux directions de l'espace formant entre elles l'angle γ. En particulier, en choisissant $\theta' = \theta$, $\phi' = \phi$, $\theta'' = \phi'' = 0$, on en déduit :

$$P_l(\cos\theta) = \frac{4\pi}{2l+1}\sum_{m=-l}^{+l} Y_{lm}(\theta, \phi)\, Y_{lm}^*(0,\, 0) \ , \tag{18.122}$$

mais comme $Y_{lm}(0,\, 0) = \delta_{m0}\sqrt{\frac{2l+1}{4\pi}}$ (voir (18.74)), il vient $P_l(\cos\theta) = \sqrt{\frac{4\pi}{2l+1}}\, Y_{l0}(\theta, \phi)$, en accord avec (18.120). Compte tenu de (18.122), le développement (18.114) peut aussi s'écrire :

$$\frac{1}{r_{12}} = \sum_{l\in\mathbb{N}}\sum_{m=-l}^{+l} \frac{4\pi}{2l+1}\frac{r_<^l}{r_>^{l+1}}\, Y_{lm}(\theta_1, \phi_1)\, Y_{lm}^*(\theta_2, \phi_2) \tag{18.123}$$

Cette forme est utile pour pouvoir jouer avec les relations d'orthonormalisation des harmoniques sphériques lorsque les états propres sont exprimés à l'aide de ces dernières.

Les expressions des premières harmoniques sphériques sont :

$$Y_{00}(\theta, \phi) = \frac{1}{\sqrt{4\pi}} \ , \tag{18.124}$$

$$Y_{10}(\theta, \phi) = \sqrt{\frac{3}{4\pi}}\cos\theta \ , \qquad Y_{1\pm1}(\theta, \phi) = \mp\sqrt{\frac{3}{8\pi}}\sin\theta\, \mathrm{e}^{\pm \mathrm{i}\phi} \ , \tag{18.125}$$

$$Y_{20}(\theta, \phi) = \sqrt{\frac{5}{16\pi}}\,(3\cos^2\theta - 1) \ , \qquad Y_{2\pm1}(\theta, \phi) = \mp\sqrt{\frac{15}{8\pi}}\sin\theta\,\cos\theta\, \mathrm{e}^{\pm \mathrm{i}\phi} \ , \tag{18.126}$$

$$Y_{2\pm2}(\theta, \phi) = \sqrt{\frac{15}{32\pi}}\sin^2\theta\, \mathrm{e}^{\pm 2\mathrm{i}\phi} \ . \tag{18.127}$$

▲

18.4 Cas particuliers des moments cinétiques $j = 1/2$ et $j = 1$

Après les traitements formels des paragraphes précédents, il est bon de revenir à des cas un peu plus concrets, d'autant plus que les deux choix de j détaillés ci-dessous sont extrêmement fréquents en pratique et jouent de ce fait un rôle important dans les applications.

Commençons par traiter le cas $j = 1/2$. Comme on le sait, il ne peut s'agir d'un moment cinétique orbital : c'est donc forcément un moment de spin et pour rappeler ceci, on désignera par $\vec{S}$ (et non pas $\vec{J}$) le moment cinétique considéré.

Les états propres de $(\vec{S}^2, S_z)$ sont notés $|s\,m\rangle$ avec $s = 1/2$, $m = \pm 1/2$. L'espace des états de spin $S = 1/2$ est ainsi de dimension $\mathrm{g}_{\frac{1}{2}} = (2s+1)_{s=1/2} = 2$. On a :

$$\vec{S}^2\,|\frac{1}{2}\,m\rangle = \frac{3}{4}\hbar^2\,|\frac{1}{2}\,m\rangle \ , \quad S_z\,|\frac{1}{2}\,m\rangle = m\hbar\,|\frac{1}{2}\,m\rangle \qquad (m = \pm\frac{1}{2}) \ . \tag{18.128}$$

Les opérateurs S_u ($u = x, y, z$) sont représentés par des matrices 2×2, qu'il est facile d'écrire. En ce qui concerne $\vec{S}^2$, c'est vite fait, tous les vecteurs propres correspondant à la même[35] valeur propre $3\hbar^2/4$:

$$\vec{S}^2 = \frac{3}{4}\hbar^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \equiv \frac{3}{4}\hbar^2\,\mathbf{1}_2 \ . \tag{18.129}$$

$\vec{S}^2$ est réellement le prototype de l'opérateur scalaire (au sens où ceci a été défini au chapitre 17) : sa matrice est proportionnelle à la matrice identité. Venons-en à S_z ; sur la base ordonnée $(|\frac{1}{2}\ +\frac{1}{2}\rangle, |\frac{1}{2}\ -\frac{1}{2}\rangle)$, il vient :

$$\boxed{S_z = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}} \tag{18.130}$$

Pour avoir les matrices de S_x et S_y il suffit de passer par les opérateurs $S_\pm$. D'après les formules générales (18.38), on a :

$$S_-|\frac{1}{2}\,\frac{1}{2}\rangle = \hbar\sqrt{\frac{1}{2}(\frac{1}{2}+1) - \frac{1}{2}(\frac{1}{2}-1)}\,|\frac{1}{2}\ -\frac{1}{2}\rangle = \hbar|\frac{1}{2}\ -\frac{1}{2}\rangle \ , \quad S_-\,|\frac{1}{2}\ -\frac{1}{2}\rangle = 0 \ . \tag{18.131}$$

Comme d'habitude, la matrice d'un opérateur se forme en portant en colonne les composantes du transformé du vecteur de la base ayant le même numéro d'ordre que la colonne considérée. D'où la matrice représentant S_- sur la base ordonnée ci-dessus :

$$S_- = \hbar \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \ . \tag{18.132}$$

S_+ est l'adjoint de S_- ; la base étant orthonormée, sa matrice est l'adjointe de celle de S_- :

$$S_+ = \hbar \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \ . \tag{18.133}$$

Il suffit maintenant d'inverser $S_\pm = S_x \pm \mathrm{i}S_y$ pour obtenir :

$$\boxed{S_x = \frac{1}{2}(S_+ + S_-) = \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad S_y = \frac{1}{2\mathrm{i}}(S_+ - S_-) = \frac{\hbar}{2}\begin{bmatrix} 0 & -\mathrm{i} \\ +\mathrm{i} & 0 \end{bmatrix}} \tag{18.134}$$

[35] Pour la simplicité des notations, et quand aucune ambiguïté n'est à craindre, on utilise le même symbole pour désigner un opérateur, et sa matrice sur une base donnée. Par ailleurs, quand cela se révèle utile, on flanque le symbole **1** d'un indice égal à la dimension de l'espace où agit cet opérateur-identité.

Pour se faire la main, il n'est pas inutile de vérifier que ces différentes matrices satisfont bien les relations de commutation caractéristiques d'un moment cinétique.

Les trois matrices 2×2 ainsi mises en évidence sont les célèbres matrices de Pauli, omniprésentes en Mécanique quantique et qu'il vaut la peine d'écrire explicitement :

$$\boxed{\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma_y = \begin{bmatrix} 0 & -\mathrm{i} \\ +\mathrm{i} & 0 \end{bmatrix} \qquad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}} \tag{18.135}$$

Un spin 1/2, $\vec{S}$, peut donc s'écrire vectoriellement :

$$\boxed{\vec{S} = \frac{\hbar}{2}\vec{\sigma}} \tag{18.136}$$

Les matrices de Pauli satisfont des relations de commutation qui ressemblent évidemment à celle d'un moment cinétique, à des facteurs près ; très précisément :

$$\boxed{[\sigma_x,\, \sigma_y] = 2\mathrm{i}\,\sigma_z \ , \qquad [\sigma_y,\, \sigma_z] = 2\mathrm{i}\,\sigma_x \ , \qquad [\sigma_z,\, \sigma_x] = 2\mathrm{i}\,\sigma_y} \tag{18.137}$$

En outre, il est facile de voir que les *anticommutateurs* des σ_u sont nuls[36] :

$$\boxed{\{\sigma_x,\, \sigma_y\} = 0 \ , \qquad \{\sigma_y,\, \sigma_z\} = 0 \ , \qquad \{\sigma_z,\, \sigma_x\} = 0} \tag{18.138}$$

de sorte que, par combinaison avec (18.137) :

$$\sigma_x\,\sigma_y = \mathrm{i}\,\sigma_z \ , \qquad \sigma_y\,\sigma_z = \mathrm{i}\,\sigma_x \ , \qquad \sigma_z\,\sigma_x = \mathrm{i}\,\sigma_y \ . \tag{18.139}$$

Les matrices de Pauli ont une trace nulle ; par ailleurs $\sigma_u^2 = \mathbf{1}_2$. Au total, avec (18.139) et $\mathrm{Tr}\mathbf{1}_2 = 2$, on a :

$$\boxed{\mathrm{Tr}(\sigma_u\sigma_v) = 2\delta_{uv}} \tag{18.140}$$

Ainsi, les matrices $M_u \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}\sigma_u$ sont orthogonales vis-à-vis du produit scalaire $(M_u,\, M_v)$ défini comme :

$$\boxed{(M_u,\, M_v) \stackrel{\text{déf}}{=} \mathrm{Tr}\,(M_u^\dagger M_v)} \tag{18.141}$$

qui a les mêmes propriétés de sesquilinéarité que le produit scalaire de deux vecteurs ; pour les matrices de Pauli, qui sont hermitiques, l'opération symbolisée par † n'est pas nécessaire. On définit aussi les $\sigma_\pm \stackrel{\text{déf}}{=} \sigma_x \pm \mathrm{i}\sigma_y$, qui satisfont :

$$[\sigma_z,\, \sigma_\pm] = \pm 2\mathrm{i}\,\sigma_\pm \ , \qquad [\sigma_+,\, \sigma_-] = 4\,\mathrm{i}\,\sigma_z \ . \tag{18.142}$$

Il existe une égalité importante, qu'il n'est pas difficile de démontrer[37]. Soit $\vec{A}$ et $\vec{B}$ deux vecteurs scalaires vis-à-vis des opérateurs de spin, au sens défini dans le chapitre 17 ;

[36] L'anticommutateur est $\{\sigma_u,\, \sigma_v\} \stackrel{\text{déf}}{=} \sigma_u\sigma_v + \sigma_v\sigma_u$.

[37] Toutes les directions de l'espace jouent ici le même rôle (isotropie galiléenne) : il suffit donc de démontrer (18.143) avec un choix particulier de l'un des vecteurs scalaires, par exemple $\vec{A} = A\,\vec{e}_x$. La distribution du commutateur et l'usage répété des relations (18.137) - (18.142) conduisent aisément à l'égalité (18.143) – voir exercice 18.7.3.

Par "vecteur scalaire", on entend ici un triplet de trois objets qui commutent avec les matrices de Pauli ; il peut s'agir d'un opérateur vectoriel n'agissant que sur les variables d'espace, des trois composantes d'un champ externe non soumis à la quantification, ou de trois scalaires au sens ordinaire.

on a alors l'égalité :

$$\boxed{(\vec{\sigma}.\vec{A})\,(\vec{\sigma}.\vec{B}) = \vec{A}.\vec{B}\,\mathbf{1}_2 + \mathrm{i}\,\vec{\sigma}.(\vec{A}\times\vec{B})} \tag{18.143}$$

Les deux vecteurs $\vec{A}$ et $\vec{B}$ peuvent être des opérateurs ne dépendant que des grandeurs dynamiques spatiales (orbitales), auquel cas il est impératif de maintenir le même ordre entre eux dans les deux membres.

Notons enfin que l'égalité $\sigma_u^2 = \mathbf{1}_2$ est l'expression la plus élémentaire du théorème de Cayley - Hamilton pour un opérateur dont le spectre est $\{-1, +1\}$:

$$\boxed{\sigma_u^2 - \mathbf{1}_2 = 0} \tag{18.144}$$

Ce théorème est très utile pour des dimensions supérieures N, puisqu'il permet d'exprimer toute fonction analytique d'opérateur sous la forme d'un *polynôme* de degré $N-1$. Barakat et Baumann [32] ont montré que les coefficients de ce polynôme peuvent se déduire des polynômes de Lucas.

Tout opérateur Ω agissant dans un espace de dimension 2 est représentable par des matrices 2×2 et a donc 4 degrés de liberté (complexes). Comme les matrices σ_u et $\mathbf{1}_2$ sont toutes linéairement indépendantes, la matrice[38] de Ω peut s'exprimer en combinaison linéaire de l'identité et des trois matrices de Pauli ; un tel développement est forcément une certaine forme linéaire :

$$\Omega = \omega_0\mathbf{1}_2 + \omega_x\sigma_x + \omega_y\sigma_y + \omega_z\sigma_z \equiv \omega_0\mathbf{1}_2 + \vec{\omega}.\vec{\sigma} \ ; \tag{18.145}$$

en utilisant maintenant (18.140), on voit que :

$$\boxed{\omega_0 = \frac{1}{2}\mathrm{Tr}\,\Omega \qquad \omega_u = \frac{1}{2}\mathrm{Tr}\,(\Omega\sigma_u)} \tag{18.146}$$

Si A désigne un opérateur hermitique, on a $A = a_0\,\mathbf{1}_2 + \vec{a}.\vec{\sigma}$ avec :

$$(a_0\mathbf{1}_2 + \vec{a}.\vec{\sigma})^\dagger = a_0\,\mathbf{1}_2 + \vec{a}.\vec{\sigma} \ ; \tag{18.147}$$

comme $\sigma_u^\dagger = \sigma_u$ et $\mathbf{1}_2^\dagger = \mathbf{1}_2$, on en déduit que a_0 et $\vec{a}$ sont tous réels :

$$\boxed{A \ \text{hermitique}: A = a_0\mathbf{1}_2 + \vec{a}.\vec{\sigma} \qquad (a_0,\, \vec{a}) \in \mathbb{R}^4} \tag{18.148}$$

Un tel opérateur est donc en correspondance biunivoque avec un certain vecteur $\vec{a}$ de l'espace $\mathbb{R}^3$; cette représentation sera fréquemment utilisée dans la suite, par exemple pour l'opérateur densité (représentation de Bloch) dans la discussion de l'effet Zénon (chapitre 21, section 21.1). Explicitement, (18.148) se lit :

$$A = \begin{bmatrix} a_0 + a_z & a_x - \mathrm{i}a_y \\ a_x + \mathrm{i}a_y & a_0 - a_z \end{bmatrix} \tag{18.149}$$

[38]Rappelons que l'ensemble des opérateurs linéaires définis dans un espace de dimension d peut être muni d'une structure d'algèbre ; l'espace vectoriel induit a pour dimension d^2.

Noter que le déterminant de A est égal à :

$$\boxed{\det A = a_0^2 - \vec{a}^{\,2}} \tag{18.150}$$

Si U est unitaire, on peut de même écrire $U = u_0\mathbf{1}_2 + \vec{u}.\vec{\sigma}$; la relation caractéristique $UU^\dagger = \mathbf{1}$ donne :

$$(u_0\mathbf{1}_2 + \vec{u}.\vec{\sigma})(u_0^*\mathbf{1}_2 + \vec{u}^*.\vec{\sigma}) = \mathbf{1}_2 \ ; \tag{18.151}$$

en développant, et en utilisant (18.143), on en déduit $u_0u_0^* + \vec{u}.\vec{u}^* = \mathbf{1}$:

$$\boxed{U\,\text{unitaire}: U = u_0\,\mathbf{1}_2 + \vec{u}.\vec{\sigma} \qquad |u_0|^2 + \|\vec{u}\|^2 = 1} \tag{18.152}$$

En particulier, si $\vec{n}$ est un vecteur unitaire ($\|\vec{n}\| = 1$), on a :

$$\boxed{\mathrm{e}^{\mathrm{i}\alpha\vec{n}.\vec{\sigma}} = \cos\alpha\,\mathbf{1}_2 + \mathrm{i}\sin\alpha\,\vec{n}.\vec{\sigma}} \tag{18.153}$$

qui constitue une sorte de formule de Moivre quantique. Ces décompositions sont très utiles, et utilisées à plusieurs reprises dans le problème 18.7.11 où est étudié en détail un phénomène important : l'oscillation de Rabi (elles sont d'ailleurs vraies quel que soit le scalaire α, réel ou complexe). Notons enfin que le développement $\Omega = \omega_0\mathbf{1}_2+\vec{\omega}.\vec{\sigma}$ peut être vu comme un cas particulier d'un résultat plus général ; en effet compte tenu de (18.144), on voit que si $f(x)$ est une fonction développable en série entière, $f(x) \stackrel{\text{déf}}{=} \sum_{n\in\mathbb{N}} c_nx^n$, alors $f(\sigma_u)$ est un polynôme du premier degré en σ_u :

$$\boxed{f(\alpha\vec{n}.\vec{\sigma}) \stackrel{\text{déf}}{=} \sum_{m\in\mathbb{N}} c_m(\alpha\vec{n}.\vec{\sigma})^m = f_0(\alpha)\mathbf{1}_2 + f_1(\alpha)\,\vec{n}.\vec{\sigma} \qquad (\|\vec{n}\| = 1)} \tag{18.154}$$

Il est remarquable que le spin 1/2 produise toujours spontanément des fonctions *linéaires* des opérateurs.

Passons maintenant au moment cinétique $j = 1$, qui peut être soit orbital, soit de spin (il ne peut résulter de la combinaison des deux, puisque le moment orbital ne peut être demi-entier et que pour avoir $j = 1$ avec deux moments cinétiques il faut que les deux soient demi-entiers – voir section 18.5). Toujours en excluant la considération d'autres degrés de liberté, la dimension de l'espace est égale à $\mathrm{g}_j = (2j + 1)_{j=1} = 3$. Les opérateurs J_x, J_y et J_z sont donc représentés par des matrices 3×3. La valeur propre de $\vec{J}^2$ est $[j(j + 1)]_{j=1}\hbar^2 = 2\hbar^2$ et le nombre magnétique m peut prendre les trois valeurs $-1, 0, +1$. Bien sûr :

$$\vec{J}^2 = \hbar^2 \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \equiv 2\hbar^2\mathbf{1}_3 \ , \tag{18.155}$$

qui est à nouveau le prototype de l'opérateur scalaire, cette fois dans l'espace à trois dimensions. En ordonnant la base suivant $|1\ +1\rangle, |1\,0\rangle, |1\ -1\rangle$, la matrice de J_z est :

$$\boxed{J_z = \hbar \begin{bmatrix} +1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}} \tag{18.156}$$

Déterminons maintenant la matrice de J_- en appliquant cet opérateur aux trois vecteurs de base :

$$J_-|1\ +1\rangle = \hbar\sqrt{1(1+1)-1(1-1)}\,|1\,0\rangle = \hbar\sqrt{2}\,|1\,0\rangle\ , \tag{18.157}$$

$$J_-|1\ +0\rangle = \hbar\sqrt{1(1+1)-0(1-1)}\,|1\,0\rangle = \hbar\sqrt{2}\,|1\ -1\rangle\ , \tag{18.158}$$

et bien sûr $J_-|1\ -1\rangle = 0$. D'où la matrice de J_-, obtenue en portant en colonne les composantes des vecteurs $J_-|1\,m\rangle$, puis celle de J_+ en prenant l'adjointe (puisque la base est orthonormée) :

$$J_- = \hbar\begin{bmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix}\ , \qquad J_+ = \hbar\begin{bmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{bmatrix}\ . \tag{18.159}$$

Les deux matrices de $J_x = \frac{1}{2}(J_+ + J_-)$ et $J_y = \frac{1}{2\mathrm{i}}(J_+ - J_-)$ s'en déduisent :

$$\boxed{J_x = \hbar\begin{bmatrix} 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 \end{bmatrix} \qquad J_y = \hbar\begin{bmatrix} 0 & -\mathrm{i}/\sqrt{2} & 0 \\ \mathrm{i}/\sqrt{2} & 0 & -\mathrm{i}/\sqrt{2} \\ 0 & \mathrm{i}/\sqrt{2} & 0 \end{bmatrix}} \tag{18.160}$$

Ces résultats donnent les transformés des vecteurs propres par les différentes composantes du moment cinétique pour $j = 1$. Comme j est entier, il peut s'agir d'un moment orbital et on peut notamment les appliquer aux harmoniques sphériques $l = 1$; on a ainsi :

$$J_x Y_{1\pm1}(\theta,\phi) = \frac{\hbar}{\sqrt{2}} Y_{10}\ , \qquad J_x Y_{10}(\theta,\phi) = \frac{\hbar}{\sqrt{2}}[Y_{11}(\theta,\phi) + Y_{1\,-1}(\theta,\phi)]\ , \tag{18.161}$$

$$J_y Y_{1\pm1}(\theta,\phi) = \pm\frac{\mathrm{i}\hbar}{\sqrt{2}} Y_{10}\ , \quad J_y Y_{10}(\theta,\phi) = \frac{\mathrm{i}\hbar}{\sqrt{2}}[-Y_{11}(\theta,\phi) + Y_{1-1}(\theta,\phi)]\ , \tag{18.162}$$

etc., relations que l'on peut aussi trouver – plus laborieusement – en utilisant les expressions (18.50). Notons aussi que le théorème de Cayley - Hamilton montre que les trois opérateurs J_u pour $j = 1$ satisfont :

$$\boxed{J_u^3 - J_u = 0} \tag{18.163}$$

puisque tous les trois ont le même spectre $\{-1,\,0,\,+1\}$, formé des solutions de l'équation $\lambda^3 - \lambda = 0$. Il en résulte que, pour $j = 1$, toute fonction de J_u développable en série entière est toujours exprimable sous la forme d'un polynôme de degré 2 en J_u :

$$\boxed{f(\alpha J_u) \stackrel{\text{déf}}{=} \sum_{n\in\mathbb{N}} c_n(\alpha) J_u^n = f_0(\alpha)\mathbf{1}_3 + f_1(\alpha) J_u + f_2(\alpha) J_u^2} \tag{18.164}$$

C'est l'équivalent dans un espace de dimension $d = 3$ de (18.154) en dimension 2. En particulier :

$$\boxed{\mathrm{e}^{\mathrm{i}\alpha J_u} = e_0(\alpha)\mathbf{1}_3 + e_1(\alpha) J_u + e_2(\alpha) J_u^2} \tag{18.165}$$

où les $e_{\mu=0,\,1,\,2}$ sont certaines fonctions de α – ce ne sont pas de simples exponentielles, voir par exemple le problème 18.7.12, éq. (18.294).

Les matrices ci-dessus peuvent être utilisée comme base pour les matrices 3×3, tout comme les matrices de Pauli σ_u pour un espace de dimension 2. Toutefois, les 4 matrices $\mathbf{1}_3$ et J_u ne suffisent pas puisqu'une matrice 3×3 quelconque a 9 degrés de liberté complexes ; il faut donc leur adjoindre d'autres matrices, linéairement indépendantes ; il est commode de les prendre toutes mutuellement orthogonales suivant le produit scalaire $(M_i,\, M_j) \stackrel{\text{déf}}{=} \mathrm{Tr}\,(M_i^\dagger M_j) = \delta_{ij}$.

18.5 Addition de deux moments cinétiques. Coefficients de Clebsch-Gordan

18.5.1 Comment apparaît la somme des moments cinétiques

L'addition (aussi appelée composition) des moments cinétiques est une opération dont l'importance se mesure sur la simple considération de quelques situations exemplaires. Ainsi, soit d'abord une particule nantie d'un moment cinétique intrinsèque $\vec{S}$ mais ayant par ailleurs un mouvement orbital lui conférant un moment orbital $\vec{L}$; il est naturel de définir un moment cinétique total $\vec{J}$ par la relation :

$$\vec{J} \stackrel{\text{déf}}{=} \vec{L} + \vec{S} \ . \tag{18.166}$$

Cette relation est naturelle mais, telle quelle, n'a pas vraiment de contenu physique et, au fond, n'a rien d'évident ; on verra au chapitre 20 pourquoi elle s'impose comme une nécessité. Quoi qu'il en soit, une fois admise (18.166), la description quantique d'une particule avec spin exige la connaissance des états propres de ce moment total : il faut donc savoir comment ils se construisent à partir des états propres de $\vec{L}$ et de ceux de $\vec{S}$.

Une problématique analogue se pose pour un système de N particules (supposées sans spin pour simplifier) possédant chacune un moment orbital ($\vec{L}_n$ pour la n^{e} particule) ; il est tout aussi naturel de définir le moment cinétique total $\vec{L}$ par la somme vectorielle :

$$\vec{L} \stackrel{\text{déf}}{=} \sum_{n=1}^{N} \vec{L}_n \ . \tag{18.167}$$

À nouveau, ceci n'a rien d'évident, quoique le contenu physique de (18.167) est transparent par référence à l'invariance euclidienne par rotation : l'opérateur de rotation pour plusieurs particules implique le produit des rotations individuelles de chacune d'entre elles, soit $\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta_n \vec{L}_n \vec{u}_n}$. Comme les opérateurs $\vec{L}_n$ agissent sur des variables différentes, les commutateurs $[\vec{L}_n\vec{u}_n,\, \vec{L}_{n'}\vec{u}_{n'}]$ sont nuls si $n \neq n'$, de sorte que le produit des exponentielles est égal à l'exponentielle de la somme :

$$\prod_{n=1}^{N} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta_n \vec{L}_n \vec{u}_n} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\sum_{n=1}^{N}\theta_n\, \vec{L}_n \vec{u}_n} \ . \tag{18.168}$$

L'invariance n'est obtenue que si toutes les particules tournent de la même façon (même angle, même axe), ce qui fait alors naturellement apparaître la somme présente au second membre de (18.167), où toutes les particules sont traitées à l'identique. Cela étant acquis, la question est à nouveau de déterminer les valeurs du moment total $\vec{L}$ connaissant celles des moments cinétiques individuels $\vec{L}_n$.

On imagine aisément que la procédure de composition est assez lourde quand il s'agit de combiner plusieurs moments cinétiques : avec N moments cinétiques $\vec{J}_n$ dont chacun a la valeur propre j_n, l'espace des états est de dimension $g_N = \prod_{n=1}^{N}(2j_n + 1)$; clairement, il suffit de savoir en additionner *deux* pour, de proche en proche, pouvoir en principe en combiner un nombre arbitraire. C'est pourquoi il suffit de développer le formalisme en se bornant à l'addition de *deux* moments cinétiques quelconques.

La toute première chose dont il faut se convaincre est que la somme de deux moments cinétiques est encore un moment cinétique. Avant d'établir ce résultat, il est utile d'affiner les notations et en particulier d'introduire à l'occasion la notion de produit (tensoriel) de deux espaces vectoriels. Pour fixer les idées, on raisonne dans le cas d'une particule possédant un moment orbital $\vec{L}$ et un spin $\vec{S}$, mais ce qui suit vaut tout autant pour deux moments cinétiques quelconques $\vec{J}_1$ et $\vec{J}_2$, quelle que soit leur origine physique.

Le moment orbital agit sur les degrés de liberté ayant un équivalent classique (coordonnées au sens usuel), le plus souvent appelées variables d'espace. S'il n'était le spin, les états quantiques de cette particule seraient des fonctions d'onde ordinaires que l'on peut toujours représenter avec les états propres $\psi_n(\vec{r})$ du Hamiltonien H. En réalité, quand la particule possède un spin (le cas le plus courant), il convient de considérer tous les états formés en associant successivement une fonction $\psi_n(\vec{r})$ donnée à chacun de tous les états de spin possibles. Pour un spin $S = 1/2$, chaque $\psi_n(\vec{r})$ permet de construire ainsi *deux* états quantiques distincts, que l'on peut noter $\psi_{n\,+}(\vec{r})$ et $\psi_{n\,-}(\vec{r})$ en introduisant juste le signe de la valeur propre $\pm\hbar/2$ de la composante S_z ; si l'on épuise tous les états les uns après les autres, on voit que le nombre total d'états quantiques "complets"[39] est le produit du nombre d'états orbitaux par le nombre d'états de spin. L'espace vectoriel correspondant est par définition le produit *tensoriel* des espaces d'espace et de spin. Dans chacun de ces deux espaces, il existe un opérateur identité, $\mathbf{1}$, que l'on peut noter $\mathbf{1}_{\text{orb}}$ pour l'espace "orbital" et $\mathbf{1}_{\text{spin}}$ pour l'espace de spin.

Maintenant, quand on combine $\vec{L}$ et $\vec{S}$ pour fabriquer $\vec{J}$, l'écriture précise correcte est de fait :

$$\vec{J} = \vec{L} \otimes \mathbf{1}_{\text{spin}} + \mathbf{1}_{\text{orb}} \otimes \vec{S} \ , \qquad (18.169)$$

où le symbole $\otimes$ est utilisé pour distinguer l'opération en question ("produit" de deux opérateurs agissant dans des espaces différents) du produit ordinaire de deux opérateurs agissant dans un même espace [4]. De la même façon, si on combine deux moments orbitaux de deux particules distinctes, l'écriture rigoureuse pour le moment total est :

$$\vec{L} = \vec{L}_1 \otimes \mathbf{1}_2 + \mathbf{1}_1 \otimes \vec{L}_2 \ , \qquad (18.170)$$

[39]au sens : donnant une description complète de la particule, avec tous ses degrés de liberté.

où $\mathbf{1}_i$ désigne ici l'opérateur identité dans l'espace des états de la particule i ($i = 1, 2$) (l'ordre des opérateurs associés dans le produit tensoriel $\otimes$ est visiblement sans importance). Ce luxe de précautions dans l'écriture est parfois nécessaire : toutefois, quand les idées sont bien en place et qu'aucune ambiguïté n'est à craindre, on s'en tient usuellement à l'écriture plus simple, du genre $\vec{L} = \vec{L}_1 + \vec{L}_2$.

Cela étant précisé, il est aisé de montrer que la somme de deux moments cinétiques est de fait un moment cinétique. Revenant pour l'instant au cas $\vec{J} = \vec{L} + \vec{S}$ afin de fixer les idées, les deux opérateurs $\vec{L}$ et $\vec{S}$ agissent dans deux espaces différents et, de ce fait, commutent entre eux :

$$[L_u,\, S_v] = 0 \quad \forall\, u,\, v \ , \tag{18.171}$$

puisque chaque opérateur ne peut modifier les degrés de liberté que l'autre affecte et réciproquement. Il résulte de (18.171) que[40] :

$$[J_u,\, J_v] = [L_u + S_u,\, L_v + S_v] = [L_u,\, L_v] + [S_u,\, S_v] = \mathrm{i}\hbar\varepsilon_{uvw}L_w + \mathrm{i}\hbar\varepsilon_{uvw}S_w \equiv \mathrm{i}\hbar\varepsilon_{uvw}J_w \tag{18.172}$$

Les mêmes arguments valent pour les moments cinétiques orbitaux de deux particules différentes numérotées n et n' : agissant sur des variables d'espace différentes ($\vec{r}_n$ et $\vec{r}_{n'}$, $n \neq n'$), ils commutent entre eux :

$$[L_{nu},\, L_{n'v}] = 0 \quad \forall\, u,\, v,\, n \neq n' \ ; \tag{18.173}$$

en conséquence :

$$\Big[\sum_n L_{nu},\, \sum_{n'} L_{n'v}\Big] = \sum_{nn'}[L_{nu},\, L_{n'v}] = \sum_n [L_{nu},\, L_{nv}] = \sum_n \mathrm{i}\hbar\varepsilon_{uvw}L_{nw} \equiv \mathrm{i}\hbar\varepsilon_{uvw}L_w \tag{18.174}$$

En définitive, on voit bien que la somme de plusieurs moments cinétiques est effectivement un moment cinétique, quels que soient leur nombre et leur nature.

18.5.2 Somme de deux moments cinétiques

Abordons maintenant à proprement parler la composition de deux moments cinétiques $\vec{J}_1$ et $\vec{J}_2$, additionnés pour former le moment total $\vec{J} = \vec{J}_1 \otimes \mathbf{1}_2 + \mathbf{1}_1 \otimes \vec{J}_2 \rightarrow \vec{J}_1 + \vec{J}_2$. Cette sous-section est quelque peu indigeste, mais ne présente pas de difficulté réelle, hormis la relative complexité des écritures.

Chacun des moments cinétiques a ses états propres individuels, $|j_1 m_1\rangle$ et $|j_2 m_2\rangle$. Comme $\vec{J}_1$ et $\vec{J}_2$ commutent, on peut trouver des vecteurs propres communs aux quatre opérateurs $\vec{J}_1^{\,2}$, J_{1z}, $\vec{J}_2^{\,2}$, J_{2z}, que l'on note $|j_1 j_2 m_1 m_2\rangle$ et qui sont le produit tensoriel des états de chaque moment cinétique :

$$|j_1 j_2 m_1 m_2\rangle \overset{\text{déf}}{=} |j_1 m_1\rangle \otimes |j_2 m_2\rangle \tag{18.175}$$

[40] ε_{uvw} est toujours le tenseur antisymétrique qui permet d'écrire commodément les relations de commutation caractéristiques d'un moment cinétique (voir (18.4)).

ce que l'on écrit plus simplement, quand aucune ambiguïté n'est à craindre :

$$|j_1 j_2 m_1 m_2\rangle = |j_1 m_1\rangle \, |j_2 m_2\rangle \ . \tag{18.176}$$

j_1 et j_2 étant fixés, il existe $d(j_1, j_2) = (2j_1 + 1)(2j_2 + 1)$ vecteurs linéairement indépendants de ce type[41], constituant une base pour un certain sous-espace $\mathcal{E}_{j_1 j_2}$. Chacun d'entre eux est propre de $(\vec{J}_1^2,\, J_{1z},\, \vec{J}_2^2,\, J_{2z})$.

À partir de maintenant, on raisonne exclusivement dans le sous-espace [4] $\mathcal{E}_{j_1 j_2}$ de dimension égale à $d(j_1, j_2)$. Il existe un autre ensemble d'observables qui commutent ; en effet, les $\vec{J}_i^2$ $(i = 1, 2)$ commutent avec $\vec{J}^2$:

$$[\vec{J}_1^2,\, \vec{J}^2] = [\vec{J}_1^2,\, \vec{J}_1^2 + \vec{J}_2^2 + 2\vec{J}_1.\vec{J}_2] = [\vec{J}_1^2,\, \vec{J}_2^2] + 2\,[\vec{J}_1^2,\, \vec{J}_1].\vec{J}_2 = 0 \ . \tag{18.177}$$

et commutent évidemment entre eux, puisqu'ils agissent sur des variables différentes (en revanche[42], le commutateur $[J_{iz},\, \vec{J}^2]$ n'est pas nul). Ainsi, les quatre opérateurs $\vec{J}^2$, J_z, $\vec{J}_1^2$ et $\vec{J}_2^2$ ont des vecteurs propres communs, que l'on note $|JMj_1j_2\rangle$. Les vecteurs $\{|JMj_1j_2\rangle\}_{-J\le M\le +J}$ et $\{|j_1j_2m_1m_2\rangle\}_{-j_i\le m_i\le +j_i}$ forment deux bases orthonormées de $\mathcal{E}_{j_1 j_2}$, puisqu'elles sont constituées de vecteurs propres d'opérateurs hermitiques. j_1 et j_2 étant fixés – on reste bien dans $\mathcal{E}_{j_1 j_2}$ –, il existe une transformation unitaire reliant ces deux jeux de vecteurs propres orthogonaux, que l'on peut écrire formellement dans un sens :

$$|JMj_1j_2\rangle = \sum_{m_1,\, m_2} |j_1j_2m_1m_2\rangle \, \langle j_1j_2m_1m_2|JMj_1j_2\rangle \ , \tag{18.178}$$

et dans l'autre :

$$|j_1j_2m_1m_2\rangle = \sum_{J,\, M} |JMj_1j_2\rangle \, \langle\, JMj_1j_2|j_1j_2m_1m_2\rangle \ . \tag{18.179}$$

Toute la question est donc de trouver les coefficients $\langle j_1j_2m_1m_2|JMj_1j_2\rangle$, qui sont appelés *coefficients de Clebsch-Gordan* (en abrégé : CG), ou encore (à des facteurs près, voir [15]) *coefficients de Wigner-3j*. Il est fréquent de les noter plus simplement[43], en sous-entendant j_1 et j_2, qui apparaissent partout et sont les valeurs maximales de $|m_1|$ et $|m_2|$; les coefficients de Clebsch-Gordan seront ainsi à l'occasion représentés par le symbole simplifié $\langle m_1m_2|JM\rangle$:

$$\boxed{\langle j_1j_2m_1m_2|JMj_1j_2\rangle \equiv \langle m_1m_2|JM\rangle} \tag{18.180}$$

ceci a toutefois l'inconvénient de faire disparaître j_1 et j_2, d'où une autre écriture commode, hybride mais plus explicite :

$$\boxed{\langle j_1j_2m_1m_2|JMj_1j_2\rangle \equiv \langle j_1j_2m_1m_2|JM\rangle} \tag{18.181}$$

[41] Chacun d'entre eux est propre d'un jeu d'observables qui commutent et donc est orthogonal à tous les autres.

[42] Cette non-commutation se vérifie trivialement. On peut aussi observer que J_{1z} est associé à une rotation qui ne change que $\vec{J}_1$; si on fait tourner $\vec{J}_1$ sans changer $\vec{J}_2$, il est bien évident que la norme de la somme $\vec{J}_1 + \vec{J}_2$ n'est pas invariante.

[43] En particulier, c'est la notation utilisée dans [15].

Cette dernière notation est d'ailleurs parfois préférable car sans ambiguïté[44] : les 2 nombres *célibataires* à droite désignent les attributs (J, M) du moment total, les 4 nombres de gauche étant ceux des deux moments additionnés. Clairement, la répétition des symboles j_1, j_2 est de toute façon inutile.

Les CG ont été calculés une fois pour toutes, et il existe des tables[45] donnant leurs valeurs, ainsi que des formules diverses pour les obtenir (la formule de Racah, par exemple) ; toutefois, dans les cas simples, il est facile et peu coûteux de les retrouver à chaque fois. Des indications seront données plus loin sur la méthode systématique qui permet de les obtenir. On peut d'ores et déjà noter les relations d'orthogonalité importantes :

$$\sum_{J,\,M} \langle j_1 j_2 m_1 m_2 | JM j_1 j_2\rangle \langle JM j_1 j_2 | j_1 j_2 m_1' m_2'\rangle = \delta_{m_1 m_1'}\delta_{m_2 m_2'} \ , \tag{18.182}$$

$$\sum_{m_1,\,m_2} \langle JM j_1 j_2 | j_1 j_2 m_1 m_2\rangle \langle j_1 j_2 m_1 m_2 | J'M' j_1 j_2\rangle = \delta_{JJ'}\delta_{MM'} \ . \tag{18.183}$$

Ces relations résultent du fait que les CG sont les coefficients d'une transformation unitaire reliant entre elles deux bases orthonormées. Dit autrement, on a immédiatement deux formes pour la relation de fermeture dans $\mathcal{E}_{j_1 j_2}$:

$$\sum_{J,\,M} |JM j_1 j_2\rangle\langle JM j_1 j_2| = \sum_{m_1,\,m_2} |j_1 j_2 m_1 m_2\rangle\langle j_1 j_2 m_1 m_2| = \mathbf{1}_{d(j_1,\,j_2)} \ . \tag{18.184}$$

Valeurs possibles de J

Avant d'en venir à l'exposé du procédé de calcul général des CG, il convient de répondre à la question suivante : j_1 et j_2 étant fixés, quelles sont les valeurs possibles de M et J donnant les valeurs propres du couple ($\vec{J}^2$, J_z) associé au moment total ? La réponse à cette question constitue le théorème fondamental d'addition énoncé ci-dessous (voir (18.203)).

En ce qui concerne les valeurs possibles de M, donnant valeur propre $M\hbar$ de J_z, la réponse est évidente ; en effet, comme $J_z = J_{1z} + J_{2z}$, on a :

$$J_z|j_1 j_2 m_1 m_2\rangle = J_{1z}|j_1 j_2 m_1 m_2\rangle + J_{2z}|j_1 j_2 m_1 m_2\rangle \ . \tag{18.185}$$

Par définition, chaque opérateur d'indice i agit exclusivement sur les nombres de même

[44] Une fois le *bra* et le *ket* remplis avec des *nombres* explicites, l'écriture de gauche dans (18.181) pourrait conduire à des ambiguïtés, car on ne saurait plus qui est qui. On verra que la convention usuelle assure que tous les CG sont réels, de sorte que le seul ennui de la notation la plus explicite est finalement sa lourdeur d'écriture.

[45] Voir par exemple [48], § 106, [15], Appendice C I. Le livre de référence en la matière est l'ouvrage d'Edmonds [33].

indice figurant dans le vecteur produit tensoriel[46]. Dans (18.185), on a à chaque fois $J_{iz}|j_1 j_2 m_1 m_2\rangle = m_i \hbar \, |j_1 j_2 m_1 m_2\rangle$, d'où :

$$J_z|j_1 j_2 m_1 m_2\rangle = (m_1 + m_2)\hbar \, |j_1 j_2 m_1 m_2\rangle \ . \tag{18.187}$$

Autrement dit, le nombre magnétique attaché à J_z, M, n'est autre que la somme :

$$\boxed{M = m_1 + m_2} \tag{18.188}$$

Inversement, un vecteur propre de J_z de valeur propre donnée $M\hbar$ est une combinaison linéaire des seuls $|j_1 j_2 m_1 m_2\rangle$ tels que $m_1 + m_2 = M$. Ceci permet déjà de voir que la valeur maximale de J est $j_1 + j_2$: en prenant $m_1 = j_1$ et $m_2 = j_2$, M est égal à $j_1 + j_2$; comme M doit varier d'une unité entre $\pm J$, bornes comprises, il doit bien exister (au moins une fois) la valeur $j_1 + j_2$ pour J. Rappelons que chaque m_i est entier ou demi-entier, selon que j_i est entier ou demi-entier ; le caractère entier *ou* demi-entier est donc fixé dès que les j_i sont fixés, ce qui est le cas dans l'analyse en cours.

Cela étant, il est bien clair qu'une valeur donnée, M, est dégénérée : M étant fixé, il existe *a priori* plusieurs façons d'assurer que la somme $m_1 + m_2$ est de fait égale à M ; si on représente le couple (m_1, m_2) par un point dans le plan de coordonnées $x = m_1$, $y = m_2$, la droite $x + y = M$ contient (en général) plusieurs points de coordonnées entières ou demi-entières ; les j_i étant fixés, chaque m_i varie par pas d'une unité.

Pour l'instant, on ne sait rien de la dégénérescence $D(J)$ des différentes valeurs de J, que l'on ne connaît pas encore et qu'il s'agit précisément de trouver – on sait seulement que $0 \leq J \leq J_{\text{max}} = j_1 + j_2$ et que, M variant au moins d'une unité, les seules valeurs envisageables sont $j_1 + j_2$, $j_1 + j_2 - 1$, $j_1 + j_2 - 2$, etc. Dans ces notations, on peut alors dire, à J fixé, qu'il existe *a priori* $D(J)$ ensembles de $(2J + 1)$ vecteurs du type $|JM\rangle$ (multiplets) dont le M varie entre $-J$ et $+J$ par sauts d'une unité. Trouver l'ensemble des valeurs de J, c'est finalement trouver toutes les valeurs de la fonction $D(J)$ pour J entier ou demi-entier positif ou nul. Ainsi par exemple, trouver $D(J_\alpha) = 0$ signifie que la valeur J_α pour J n'existe pas avec les valeurs de (j_1, j_2) considérées. En définitive, $D(J)$ est la fonction indicatrice du domaine des valeurs possibles de J pour un couple (j_1, j_2) donné.

Pour savoir ce qu'il en est et arriver au résultat cherché, différentes démonstrations peuvent être suivies ([15], XIII.25, [25], X.C). Donnons successivement deux arguments, l'un rapide, l'autre plus détaillé.

Argument rapide Cet argument s'appuie simplement sur la dimension $d(j_1, j_2)$ de l'espace $\mathcal{E}_{j_1 j_2}$, égale à $(2j_1 + 1)(2j_2 + 1)$. C'est aussi le nombre de vecteurs $|JMj_1j_2\rangle$, J

[46] Plus généralement, un opérateur A_1 agissant seulement sur la première particule produit une combinaison linéaire où les nombres quantiques de la deuxième n'ont pas changé :

$$A_1|j_1 j_2 m_1 m_2\rangle = \sum_{j_1' m_1'} A_{1\, j_1' m_1' \, ; \, j_1 m_1} \, |j_1' j_2 m_1' m_2\rangle \ , \tag{18.186}$$

d'où, si c'est nécessaire, la notation plus précise $A_1 \otimes \mathbf{1}_2$, pour un tel opérateur, $\mathbf{1}_2$ étant l'identité pour la particule 2.

et M prenant de fait toutes leurs valeurs possibles ; $D(J)$ désignant toujours la dégénérescence de la valeur J et le nombre M pouvant, à J fixé, prendre les $g_J = 2J+1$ valeurs $-J$, $-J+1$, ..., $J-1$, J, on doit avoir impérativement :

$$\boxed{\sum_J (2J+1)D(J) = (2j_1+1)(2j_2+1)} \tag{18.189}$$

étant entendu que, d'après ci-dessus, les valeurs *a priori* possibles de J sont connues ; en convenant d'appeler j_1 le plus petit de j_1 et j_2 ($j_1 \leq j_2$), ces valeurs sont de la forme $j_2 - j_1 + k$, où l'entier k varie de 0 à $2j_1$. La condition (18.189) s'écrit ainsi plus précisément :

$$\sum_{k=0}^{2j_1} [2(j_2 - j_1 + k) + 1]D(J) = (2j_1+1)(2j_2+1) \ . \tag{18.190}$$

Supposons que $D(J) = 1$ quel que soit J, ce qui signifie que chacune de toutes les valeurs $|j_2 - j_1|$, $|j_2 - j_1| + 1$, ..., $j_1 + j_2$ est possible une fois et une seule. Sous cette hypothèse, la somme au premier membre de (18.190) s'évalue sans difficulté, et on trouve précisément qu'elle vaut $(2j_1+1)(2j_2+1)$; en d'autres termes, avec $D(J) = 1\,\forall\, J$, la condition (18.189) est satisfaite.

Supposons maintenant au contraire que l'une des valeurs de J entre $j_2 \pm j_1$ est impossible, soit $J_\emptyset$, ce qui signifie par hypothèse que $D(J_\emptyset) = 0$, et qu'il manque donc pour l'instant les $2J_\emptyset + 1$ vecteurs $|J_\emptyset\, M\rangle$ avec $-J_\emptyset \leq M \leq +J_\emptyset$. On doit retrouver ailleurs dans $\mathcal{E}_{j_1 j_2}$ autant de directions en tant que composantes $|J'_\emptyset\, M\rangle$, la valeur $J'_\emptyset$ ayant alors une dégénérescence $D(J'_\emptyset) = 2$ – mais ceci est clairement impossible avec $J'_\emptyset \neq J_\emptyset$ ($2J'_\emptyset + 1 \neq 2J_\emptyset + 1$).

Ainsi, avec $D(J_\emptyset) = 0$ et $D(J \neq J_\emptyset) \neq 0$, il faut qu'il existe au moins deux autres valeurs $J'_\emptyset$ et $J''_\emptyset$ avec une dégénérescence égale à 2 apportant ensemble $2J'_\emptyset + 1 + 2J''_\emptyset + 1$ composantes supplémentaires au total pour combler les $2J_\emptyset + 1$ directions manquantes[47] ; il faut donc :

$$2J'_\emptyset + 1 + 2J''_\emptyset + 1 = 2J_\emptyset + 1 \ ; \tag{18.191}$$

notant $J'_\emptyset$ le plus petit de $J'_\emptyset$ et $J''_\emptyset$, on doit clairement avoir $J'_\emptyset < J_\emptyset < J''_\emptyset$. Maintenant, de deux choses l'une :

- ou bien j_1 et j_2 sont tous deux entiers ou demi-entiers, auquel cas toutes les valeurs de J sont entières. Dans ces conditions, l'égalité (18.191) ne peut être satisfaite, puisque le premier membre est pair, alors que le second est impair.
- l'un des deux j_i est demi-entier et l'autre est entier ; alors, toutes les valeurs de J sont demi-entières. À nouveau, (18.191) n'est jamais vraie, puisque ses deux membres sont encore de parités différentes.

[47] L'affirmation qu'il manque $2J_\emptyset + 1$ composantes est relative au bilan prenant déjà en compte les $2J'_\emptyset + 1 + 2J''_\emptyset + 1$ composantes associées aux valeurs $J'_\emptyset$ et $J''_\emptyset$, et déjà présentes de toute façon suivant l'hypothèse où *seule* la valeur $J_\emptyset$ est absente.

Ainsi, quand l'une des valeurs escomptables de J est manquante, il est impossible de satisfaire l'inévitable condition (18.189). En supposant ensuite que deux valeurs manquent à l'appel, la même argumentation montre qu'il est encore impossible de satisfaire cette condition, et ainsi de suite. La conclusion est donc que chacune des $2\inf(j_1, j_2)+1$ valeurs prévues *a priori* pour J est de fait réalisée une fois et une seule.

Argument plus détaillé À nouveau, l'idée à garder en tête est que chaque valeur possible de J arrive en bloc avec ses $g_J = 2J+1$ composantes $M = -J, -J+1, \dots, +J$. Afin de simplifier les écritures ultérieures, posons :

$$\boxed{R_\pm(J,\, M) \stackrel{\text{déf}}{=} \sqrt{J(J+1) - M(M\pm 1)}} \tag{18.192}$$

Les relations fondamentales (18.38) s'écrivent alors $J_\pm|j\,m\rangle = \hbar R_\pm(j,\, m)\ |j\,m\pm 1\rangle$; noter aussi que l'on a :

$$\boxed{R_\pm(J,\, M\mp 1) = R_\mp(J,\, M)} \tag{18.193}$$

et $R_-(J,\, J) = \sqrt{2J}$, $R_+(J,\, J) = 0$.

Il est facile de trouver le vecteur correspondant aux valeurs extrêmes de J et M, respectivement $J_{\max}$ et $M_{\max}$, toutes deux égales à j_1+j_2, puisqu'il n'y a qu'un seul *ket* dans $\mathcal{E}_{j_1 j_2}$ ayant ces deux valeurs de J et M ; d'où l'égalité :

$$|J = j_1+j_2\, M = j_1+j_2\, j_1\, j_2\rangle = |j_1\, j_2\, m_1 = j_1\, m_2 = j_2\rangle\ . \tag{18.194}$$

Appliquons l'opérateur J_- au premier membre :

$$J_-|J = j_1+j_2\, M = j_1+j_2\, j_1\, j_2\rangle = \hbar R_-(J\,M)|J = j_1+j_2\, M-1\, j_1\, j_2\rangle\ , \tag{18.195}$$

soit :

$$J_-|J = j_1+j_2\, M = j_1+j_2\, j_1\, j_2\rangle = \hbar\sqrt{2J}|J = j_1+j_2\, M-1\, j_1\, j_2\rangle\ , \tag{18.196}$$

Comme $J_- = J_{1-} + J_{2-}$, appliquons $J_{1-} + J_{2-}$ au second membre de (18.194) :

$$(J_{1-}+J_{2-})|j_1\, j_2\, m_1 = j_1\, m_2 = j_2\rangle = \hbar R_-(j_1\, m_1)|j_1\, j_2\, m_1-1\, m_2\rangle + \hbar R_-(j_2\, m_2)|j_1\, j_2\, m_1\, m_2-1\rangle\ , \tag{18.197}$$

soit :

$$(J_{1-}+J_{2-})|j_1\, j_2\, m_1 = j_1\, m_2 = j_2\rangle = \hbar\sqrt{2j_1}|j_1\, j_2\, m_1-1\, m_2\rangle + \hbar\sqrt{2j_2}|j_1\, j_2\, m_1\, m_2-1\rangle\ . \tag{18.198}$$

En rapprochant (18.196) et (18.198), il vient :

$$\sqrt{J}|J = j_1+j_2\, M = j_1+j_2-1\, j_1\, j_2\rangle = \sqrt{j_1}|j_1\, j_2\, m_1-1\, m_2\rangle + \sqrt{j_2}|j_1\, j_2\, m_1\, m_2-1\rangle\ . \tag{18.199}$$

Cette équation donne donc la composante $M = J-1$ du vecteur ayant la valeur maximale pour J, qui apparaît comme la combinaison linéaire des deux vecteurs (connus) figurant au second membre de (18.199). Dans le sous-espace de dimension 2, où tous les vecteurs ont le même $M = j_1 + j_2 - 1$, engendré par ces deux vecteurs, on peut fabriquer la combinaison orthogonale[48] :

$$\sqrt{j_2}|j_1\, j_2\, m_1 - 1\, m_2\rangle - \sqrt{j_1}|j_1\, j_2\, m_1\, m_2 - 1\rangle \ . \tag{18.200}$$

Ce vecteur correspond forcément à une autre valeur de J, mais son M vaut $J_{\text{max}} - 1$: c'est donc la composante M maximum du vecteur ayant $J = J_{\text{max}} - 1 = j_1 + j_2 - 1$; d'où l'identification :

$$|J = j_1 + j_2 - 1\, M = j_1 + j_2 - 1\, j_1\, j_2\rangle = \sqrt{2j_2}|j_1\, j_2\, m_1 - 1\, m_2\rangle - \sqrt{2j_1}|j_1\, j_2\, m_1\, m_2 - 1\rangle \ . \tag{18.201}$$

Il suffit de recommencer le même travail sur l'équation (18.199) : on applique J_- au premier membre, ce qui produit $|J = j_1 + j_2\, M = j_1 + j_2 - 2\, j_1\, j_2\rangle$, à un facteur près. Appliquons $J_{1-} + J_{2-}$ au second membre de (18.199) : on trouve une combinaison linéaire de trois vecteurs seulement, soit $|j_1\, j_2\, m_1 - 2\, m_2\rangle$, $|j_1\, j_2\, m_1 - 1\, m_2 - 1\rangle$ et $|j_1\, j_2\, m_1\, m_2 - 2\rangle$. Le sous-espace est donc de dimension égale à 3, tous les vecteurs correspondant à un M total égal à $j_1 + j_2 - 2$; une direction de ce sous-espace est la composante $M = J - 2$ du J maximum, une deuxième dimension est occupée par la composante $M = J - 1$ du $J = J_{\text{max}} - 1$. La dernière direction disponible est donc la composante (maximum) d'un état ayant $J = J_{\text{max}} - 2$.

De proche en proche, on montre ainsi que toutes les valeurs possibles de J sont effectivement présentes. Pour chaque valeur J, il reste à chaque fois une seule direction disponible en dehors des directions déjà occupées par des composantes M appartenant à des multiplets $J' > J$ rencontrés auparavant, montrant qu'il existe un seul multiplet de $2J + 1$ composantes pour chacune des valeurs de J.

Au total, d'une façon ou d'une autre, on voit que dans tous les cas :

$$\boxed{D(J) = \begin{cases} 0 & \text{si } J < |j_1 - j_2| \text{ ou si } J > j_1 + j_2 \\ 1 & \text{si } |j_1 - j_2| \leq J \leq j_1 + j_2 \end{cases}} \tag{18.202}$$

D'où le théorème fondamental d'addition de deux moments cinétiques :

- dans le sous-espace $\{|j_1 j_2 m_1 m_2\rangle\}$ de dimension $d(j_1, j_2) = (2j_1 + 1)(2j_2 + 1)$, les valeurs possibles de J sont :

$$J = |j_1 - j_2|,\ |j_1 - j_2| + 1,\ |j_1 - j_2| + 2,\ \ldots,\ j_1 + j_2 - 1,\ j_1 + j_2 \tag{18.203}$$

[48] Rappelons que deux vecteurs propres associés à deux valeurs propres distinctes d'un même opérateur hermitique (ici $\vec{J}^2$) sont orthogonaux. Par ailleurs, on ne se soucie pas pour l'instant de normaliser les différents *kets* qui apparaissent les uns après les autres.

Ce résultat peut se résumer en disant que le nombre J doit satisfaire l'inégalité triangulaire :

$$\boxed{|j_2 - j_1| \leq J \leq j_1 + j_2} \tag{18.204}$$

qui, dans la limite des très grands nombres quantiques, n'est autre que l'inégalité de géométrie élémentaire concernant les longueurs respectives des côtés d'un triangle ($J \gg 1 \Longrightarrow \|\vec{J}\| \simeq \hbar J$).

- pour chaque valeur de J, il existe un et un seul ensemble (*multiplet*) de $(2J+1)$ vecteurs du type $|JM\rangle$, $M = -J, -J+1, \ldots, J-1, J$, propres de $(\vec{J}^2, J_z)$.

♦ *Remarques*

1. Pour la somme d'un nombre *quelconque* de moments cinétiques, $\vec{J} \stackrel{\text{déf}}{=} \sum_{i=1}^{N} \vec{J_i}$, le théorème fondamental permet d'affirmer que la valeur *maximale* de J, J_{max}, est égale à $\sum_{i=1}^{N} J_i$. Il en résulte en particulier que pour N moments cinétiques identiques $J_i = 1/2$, $J_{\text{max}} = N/2$: ainsi, la réunion des $\vec{J_i}$ est une particule (composite) de moment cinétique *entier* (boson) si N est pair, *demi-entier* (fermion) si N est impair.

2. Selon le théorème fondamental, la somme de *deux* moments cinétiques prend des valeurs équidistantes, chacune de celles-ci étant possible *une fois et une seule.* Cette dernière propriété est très particulière et disparaît quand on additionne plus de deux moments cinétiques.

   Par exemple, la somme de trois spins 1/2 fournit un état quadruplet ($S = 3/2$) et *deux* états doublets ($S = 1/2$). On retrouve bien ainsi toutes les dimensions disponibles : l'espace d'états de trois spins $S = 1/2$ a la dimension $2^3 = 8$; le quadruplet possède $g_{3/2} = (2 \times \frac{3}{2} + 1) = 4$ composantes, chaque doublet a $g_{1/2} = 2$ composantes, et au total (un quadruplet et deux doublets), la dimension de l'espace des états propres du spin total est $g_{3/2} + 2g_{1/2} = 8$ (pour le cas général de N spins 1/2, voir le problème 18.7.8).

3. Bien sûr, outre les nombres quantiques relatifs aux moments cinétiques, chaque état est presque toujours caractérisé par un ensemble α de valeurs propres d'observables qui, avec les moments cinétiques, constituent un ECOC. En pareil cas, la notation des états doit être précisée : $|\alpha_1\alpha_2, j_1j_2m_1m_2\rangle \equiv |\alpha_1, j_1m_1\rangle \otimes |\alpha_2, j_2m_2\rangle$; l'équation (18.178) s'écrit alors ($\alpha \equiv (\alpha_1, \alpha_2)$) :

$$|\alpha\, JMj_1j_2\rangle = \sum_{m_1m_2} |\alpha\, j_1j_2m_1m_2\rangle\, \langle j_1j_2m_1m_2|JMj_1j_2\rangle \ . \tag{18.205}$$

Les coefficients du développement de droite ne dépendent pas des α_i. En effet, les opérateurs des moments cinétiques, sur les bases $|JMm_1m_2\rangle$ ou $|j_1j_2m_1m_2\rangle$, sont représentés par des matrices indépendantes des autres nombres quantiques α_i ; ceci traduit le caractère purement géométrique de la composition des moments cinétiques, qui ne dépend que de la valeur des moments cinétiques additionnés, en

aucune façon des valeurs propres des autres observables appartenant à l'ECOC. Il en va de même de leurs valeurs propres et de leurs vecteurs propres, d'où l'indépendance des coefficients du développement (18.205) par rapport à l'ensemble des autres nombres quantiques α. ♦

18.5.3 Propriétés des coefficients de Clebsch - Gordan. Principes de leur méthode de calcul

Outre les propriétés d'orthogonalité déjà écrites, (18.183) et (18.182), les CG possèdent des propriétés importantes qu'il faut retenir. L'une des plus saillantes est que le coefficient $\langle j_1 j_2 m_1 m_2 | JM\rangle$ est nul si les deux conditions suivantes ne sont pas simultanément satisfaites[49] :

$$\boxed{|j_1 - j_2| \leq J \leq j_1 + j_2 \qquad m_1 + m_2 = M} \tag{18.207}$$

La première condition exprime le théorème fondamental d'addition ; la deuxième égalité est une évidence résultant de $J_z = J_{1z} + J_{2z}$.

Les CG satisfont d'importantes relations de récurrence, que l'on peut établir comme suit. Partons de (18.178) en adoptant la notation simplifiée (18.181), et appliquons $J_\pm = (J_{1\pm} + J_{2\pm})$ aux deux membres ; le premier est $R_\pm(J, M)|JM \pm 1\rangle$, par définition (voir (18.192)). Le second membre est :

$$\sum_{m_1,\, m_2} [J_{1\pm}|j_1 j_2 m_1 m_2\rangle + J_{2\pm}|j_1 j_2 m_1 m_2\rangle]\langle j_1 j_2 m_1 m_2 | JM\rangle \ , \tag{18.208}$$

soit :

$$\hbar \sum_{m_1,\, m_2} [R_\pm|j_1 j_2 m_1 \pm 1 m_2\rangle + R_\pm|j_1 j_2 m_1 m_2 \pm 1\rangle]\langle j_1 j_2 m_1 m_2 | JM\rangle \ , \tag{18.209}$$

d'où l'égalité :

$$R_\pm(J,\, M)|JM \pm 1\rangle = \sum_{m_1,\, m_2} [R_\pm(j_1,\, m_1)|j_1 j_2 m_1 \pm 1 m_2\rangle + R_\pm(j_2,\, m_2)|j_1 j_2 m_1 m_2 \pm 1\rangle]\langle j_1 j_2 m_1 m_2 | JM\rangle \ ; \tag{18.210}$$

multipliant maintenant scalairement par $\langle j_1 j_2 m'_1 m'_2 |$:

$$R_\pm(J,\, M)\langle j_1 j_2 m'_1 m'_2 | JM \pm 1\rangle = R_\pm(j_1, m'_1 \mp 1)\langle j_1 j_2 m'_1 \mp 1 m_2 | JM\rangle + R_\pm(j_2,\, m'_2 \mp 1)\langle j_1 j_2 m'_1 m_2 \mp 1 | JM\rangle \ , \tag{18.211}$$

d'où, compte tenu de (18.193) :

$$R_\pm(J,\, M)\langle j_1 j_2 m_1 m_2 | JM \pm 1\rangle = R_\mp(j_1, m_1)\langle j_1 j_2 m_1 \mp 1 m_2 | JM\rangle + R_\mp(j_2,\, m_2)\langle j_1 j_2 m_1 m_2 \mp 1 | JM\rangle \ . \tag{18.212}$$

[49]Même quand ces conditions sont satisfaites, il peut arriver que le CG soit nul. Par exemple

$$\langle j100 | J = j\, M = 0\rangle = 0 \ , \tag{18.206}$$

alors que $0 + 0 = 0$ et $j - 1 \leq j \leq j + 1$ (voir exercice 18.7.5).

Il importe de se souvenir que ces relations sont à J, j_1 et j_2 fixés. Compte tenu de la relative complexité algébrique de ces relations de récurrence, il est utile de les voir à l'œuvre explicitement.

Repartons de (18.212) avec les signes inférieurs et en prenant $M = J$, $m_1 = j_1$. Pour obtenir une relation non-triviale (pas du genre $0 = 0$), il faut, en vertu de la règle : "$M = m_1 + m_2$ sinon le CG est nul", choisir $j_1 + m_2 + 1 = M = J$, soit $m_2 = J - j_1 - 1$; comme $R_-(J, J) = \sqrt{2J}$ et $R_+(j_1, j_1) = 0$, il vient :

$$\sqrt{2J}\langle j_1 j_2 j_1 J - j_1 - 1 | J J - 1\rangle = R_+(j_2, J - j_1 - 1)\langle j_1 j_2 j_1 J - j_1 | J J\rangle \ . \qquad (18.213)$$

Si donc on connaît $\langle j_1 j_2 j_1\, J - j_1 | J\, J\rangle$, cette relation permet de calculer le coefficient CG $\langle j_1 j_2 j_1\, J - j_1 - 1 | J\, J - 1\rangle$. Repartons encore de (18.212) avec cette fois les signes supérieurs et les choix $M = J - 1$, $m_1 = j_1$, $m_2 = J - j_1$; il vient :

$$\sqrt{2J}\langle j_1 j_2 j_1 J - j_1 | J J\rangle = \sqrt{2j_1}\langle j_1 j_2 j_1 - 1 J - j_1 | J J - 1\rangle + \\ R_-(j_2, J - j_1)\langle j_1 j_2 j_1 J - j_1 - 1 | J J - 1\rangle \ . \qquad (18.214)$$

$\langle j_1 j_2 j_1 J - j_1 - 1 | J J - 1\rangle$ est obtenu de la première étape si l'on connaît $\langle j_1 j_2 j_1 J - j_1 | J J\rangle$; dès que ce dernier coefficient est connu, on peut alors trouver $\langle j_1 j_2 j_1 - 1 J - j_1 | J J - 1\rangle$, et ainsi de suite. Ainsi, pour J, j_1 et j_2 fixés, tous les CG peuvent bien se déduire de l'un d'entre eux, choisi ici comme étant $\langle j_1 j_2 j_1 J - j_1 | J J\rangle$, conformément à l'usage. Ce coefficient est non nul si $J - j_1$ est compris entre $-j_2$ et $+j_2$:

$$-j_2 \leq J - j_1 \leq j_2 \quad \Longleftrightarrow \quad j_1 - j_2 \leq J \leq j_1 + j_2 \ . \qquad (18.215)$$

Par les mêmes procédés, on peut bien évidemment exprimer tous les CG à partir de $\langle j_1 j_2 J - j_2 j_2 | J\, J\rangle$ et pour que la solution ne soit pas triviale, il faut :

$$-j_1 \leq J - j_2 \leq j_1 \quad \Longleftrightarrow \quad j_2 - j_1 \leq J \leq j_1 + j_2 \ . \qquad (18.216)$$

Au total, j_1 et j_2 étant donnés, il n'y a de solution que si l'inégalité triangulaire (18.204) est satisfaite, ce que l'on savait déjà. Elle donne l'intervalle auquel appartient le nombre J associé à la longueur du moment cinétique total, j_1 et j_2 étant donnés. Dans cet intervalle, J ne varie pas continûment : tous les entiers M, m_1 et m_2 variant d'une unité dans leurs intervalles respectifs $\pm J$, $\pm j_1$ et $\pm j_2$, le nombre J varie de même. Au total, on retrouve bien les valeurs possibles pour le moment total énoncées par le théorème d'addition (18.207). La méthode purement algébrique développée ci-dessus peut être considérée comme une démonstration alternative de ce théorème.

Par ailleurs, les CG sont des produits scalaires de vecteurs propres ; ces derniers étant définis à une phase près, il en va de même des CG. Comme on l'a vu, tous les CG peuvent se déduire de l'un d'entre eux, de sorte que l'arbitraire de phase n'existe que pour celui qui sert de point de départ à la déduction par récurrence de tous autres. La convention habituelle consiste à choisir ([4], p. 392 - 393, éq. (16.72)) :

$$\boxed{\langle j_1\, j_2\, m_1 = j_1\, m_2 = J - j_1 | J\, J\rangle \ \text{ réel et positif}} \qquad (18.217)$$

Avec cette convention, tous les CG sont *réels* – mais pas forcément tous positifs, évidemment[50] – puisque la récurrence résulte de l'application des formules générales (18.38), où tous les coefficients sont réels. Il en résulte que, dans les relations d'orthogonalité (18.182) et (18.183), on peut échanger le *bra* et le *ket* d'un produit scalaire, si nécessaire :

$$\boxed{\langle j_1\, j_2\, m_1\, m_2 | J\, M\rangle \;=\; \langle J\, M | j_1\, j_2\, m_1\, m_2\rangle} \qquad (18.218)$$

Enfin, les CG satisfont les relations suivantes ([25], $\mathrm{B_X}$ 3.) :

$$\langle j_1 j_2 m_1 m_2 | JM\rangle = (-1)^{j_1+j_2-J}\langle j_2 j_1 m_2 m_1 | JM\rangle \qquad (18.219)$$

Il peut sembler surprenant que l'ordre des deux moments cinétiques joue un rôle, étant donné que l'addition $\vec{J_1} + \vec{J_2}$ est commutative. En fait, ceci vient de la convention adoptée, (18.217), qui vient briser la symétrie intrinsèque dans l'échange $(1 \leftrightarrow 2)$: il est bien clair que les deux kets $|J J\rangle$ peuvent ne pas avoir la même phase selon que l'on choisit $\langle j_1\, j_2\, j_1\, J - j_1 | J\, J\rangle > 0$ ou $\langle j_2\, j_1\, j_2\, J - j_2 | J\, J\rangle > 0$. On a aussi :

$$\langle j_1 j_2 m_1 m_2 | JM\rangle = (-1)^{j_1+j_2-J}\langle j_1 j_2 - m_1 - m_2 | J - M\rangle\ . \qquad (18.220)$$

Pour terminer, donnons quelques indications sur la façon d'obtenir effectivement et directement tous les CG dans les cas simples (petites valeurs de j_1 et j_2). La procédure standard s'initie par la détermination des $2J+1$ composantes constituant le multiplet ayant la valeur maximale de J, soit $J = j_1 + j_2$, et en commençant par la composante $M = +J$ $(= j_1 + j_2)$, que l'on peut trouver immédiatement. En effet, dans le sous-espace $\mathcal{E}_{\alpha\, j_1 j_2}$, il existe un seul vecteur ayant cette valeur de J et cette valeur de M ; il s'agit donc d'un vecteur appartenant à la fois à $\{|j_1 j_2 m_1 m_2\rangle\}$ et à $\{|JM j_1 j_2\rangle\}$. D'où l'identité :

$$|J = j_1 + j_2\, M = j_1 + j_2\, j_1\, j_2\rangle = |j_1\, j_2\, m_1 = j_1\, m_2 = j_2\rangle\ ; \qquad (18.221)$$

appliquant maintenant J_- aux deux membres, il vient :

$$J_-|J = j_1 + j_2\, M = j_1 + j_2\, j_1\, j_2\rangle = (J_{1-} + J_{2-})|j_1\, j_2\, m_1 = j_1\, m_2 = j_2\rangle\ . \qquad (18.222)$$

Le premier membre est :

$$\hbar\sqrt{2(j_1 + j_2)}\ |J = j_1 + j_2\, M = j_1 + j_2 - 1\, j_1\, j_2\rangle\ . \qquad (18.223)$$

Le second membre est :

$$\hbar\sqrt{2}\left[\sqrt{j_1}|j_1\, j_2\, m_1 = j_1 - 1\, m_2 = j_2\rangle + \sqrt{j_2}|j_1\, j_2\, m_1 = j_1\, m_2 = j_2 - 1\rangle\right]\ , \qquad (18.224)$$

d'où :

$$\sqrt{j_1 + j_2}\,|J = j_1 + j_2\, M = j_1 + j_2 - 1\, j_1\, j_2\rangle = \\ \sqrt{j_1}\,|j_1\, j_2\, m_1 = j_1 - 1\, m_2 = j_2\rangle + \sqrt{j_2}\,|j_1\, j_2\, m_1 = j_1\, m_2 = j_2 - 1\rangle\ . \qquad (18.225)$$

[50]Les relations d'orthogonalité (18.183) et (18.182) ne sauraient être satisfaites avec des CG tous de même signe.

Sur ce développement, on lit les CG correspondants, $\langle j_1 j_2 m_1 m_2 | JM \rangle$:

$$\langle j_1\, j_2\, j_1 - 1\, j_2 | j_1 + j_2\, j_1 + j_2 - 1 \rangle = \sqrt{\frac{j_1}{j_1 + j_2}} \ . \tag{18.226}$$

$$\langle j_1\, j_2\, j_1\, j_2 - 1 | j_1 + j_2\, j_1 + j_2 - 1 \rangle = \sqrt{\frac{j_2}{j_1 + j_2}} \ . \tag{18.227}$$

Une fois trouvé le vecteur $|J = j_1 + j_2\, M = j_1 + j_2 - 1\, j_1\, j_2\rangle$, on peut obtenir la composante suivante $|J = j_1 + j_2\, M = j_1 + j_2 - 2\, j_1\, j_2\rangle$ par une nouvelle application de J_- ; la lecture du développement fournit deux nouveaux CG, et ainsi de suite pour toutes les autres composantes du multiplet $J = j_1 + j_2$.

Cela fait, on passe à la valeur immédiatement inférieure de J, soit $J = j_1 + j_2 - 1$, et on cherche sa composante maximale $M = J = j_1 + j_2 - 1$. Il y a en réalité deux vecteurs (et deux seulement) ayant cette valeur de M, car il y a deux façons seulement de choisir m_1 et m_2 pour que leur somme soit égale à $j_1 + j_2 - 1$. L'un de ces vecteurs a déjà été trouvé : c'est la deuxième composante du multiplet $J = j_1 + j_2$, ayant précisément $M = j_1 + j_2 - 1$. Le deuxième vecteur est celui que l'on cherche maintenant : il se trouve par orthogonalité avec le précédent – à une phase près, fixée finalement par la convention énoncée plus haut. Il reste maintenant à effectuer une application répétée de J_- sur le vecteur qui vient d'être trouvé pour obtenir successivement toutes les $2(j_1 + j_2 - 1) + 1$ composantes du multiplet $j_1 + j_2 - 1$. Le développement de chaque composante ainsi trouvée sur les $\{|j_1 j_2 m_1 m_2\rangle\}$ fournit de nouveaux coefficients de Clebsch-Gordan. Puis on recommence avec le multiplet d'ordre immédiatement inférieur, $J = j_1 + j_2 - 2$, et ainsi de suite.

18.6 Théorème de Wigner-Eckart. Règles de sélection

Les opérateurs scalaires et vectoriels ont été définis au chapitre 17 en relation avec leur mode de transformation par rotation ; les opérateurs scalaires sont invariants (tout comme leurs moyennes), alors que les moyennes des composantes V_u d'un opérateur vectoriel se transforment comme les trois composantes cartésiennes d'un vecteur de la géométrie élémentaire dans $\mathbb{R}^3$ (voir (17.117)).

D'un autre côté, on a vu que toute rotation géométrique $R_{\vec{u}\,\theta}$ induit dans l'espace des états une transformation unitaire $\mathcal{R}_{\vec{u}\,\theta}$ qui s'exprime à l'aide du moment cinétique selon (17.98). Ceci entraîne, en particulier, qu'un vecteur $|jm\rangle$ se transforme suivant :

$$|jm\rangle' \stackrel{\text{déf}}{=} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,\vec{u}.\vec{J}}\,|jm\rangle = \sum_{m'} |jm'\rangle\,\langle jm'|\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,\vec{u}.\vec{J}}|jm\rangle \ . \tag{18.228}$$

Le point important est que la rotation ne peut changer que les nombres quantiques magnétiques : elle ne peut modifier j, puisque les trois composantes du moment cinétique

commutent avec le carré du module et que l'opérateur de rotation ne dépend que de ces trois composantes ; il en résulte que $|jm\rangle'$ n'a de composantes non nulles que sur des *kets* de même j et donc que la somme court seulement sur[51] m'. Il est alors possible d'écrire la représentation $\mathcal{R}^{(j)}_{mm'}$ de la transformation unitaire $\mathcal{R}^{(j)}$ induite dans l'espace des états par la rotation R dans $\mathbb{R}^3$:

$$|jm\rangle' = \sum_{m'} |jm'\rangle\, R^{(j)}_{m'm} \ , \qquad R^{(j)}_{m'm} = \langle jm'|\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\, \vec{u}.\vec{J}}|jm\rangle \ , \tag{18.231}$$

et met à nouveau en évidence une matrice, dite de rotation pour faire court, qui peut maintenant avoir une dimension quelconque (entière !), égale à $2j+1$. Avec cette notation, les matrices des rotations ordinaires (et celles qui donnent le mode de transformation des opérateurs vectoriels) peuvent se noter désormais $\mathcal{R}^{(1)}_{m'm}$. En redéfinissant convenablement les composantes standard d'un opérateur vectoriel à partir de ses composantes cartésiennes (voir (17.119)), on peut toujours écrire la combinaison linéaire suivante :

$$(V^{(1)}_m)' = \mathcal{R}^{(1)} V^{(1)}_m \mathcal{R}^{(1)\dagger} = \sum_{m'} V^{(1)}_{m'}\, \mathcal{R}^{(1)}_{m'm} \ , \qquad \mathcal{R}^{(1)}_{m'm} = \langle 1m'|\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\, \vec{u}.\vec{J}}|1m\rangle \ . \tag{18.232}$$

On vérifie sans peine que $\vec{r}$ et $\vec{p}$ sont des opérateurs vectoriels. Pour une particule sans spin, le moment cinétique orbital $\vec{L} = \vec{r} \times \vec{p}$ est également un opérateur vectoriel. Pour une particule avec spin, $\vec{S}$ est un opérateur vectoriel vis-à-vis du moment cinétique *total*[52] $\vec{J} = \vec{L} + \vec{S}$. D'un autre côté, un champ externe (électrique, magnétique, ...) n'est pas un opérateur vectoriel, et pour cause : ce n'est pas un opérateur du tout ! En effet, un tel champ est supposé donné une fois pour toutes et n'appartient pas au système quantique ; toutes ses composantes commutent avec tous les opérateurs, ce sont des scalaires au sens le plus strict. Un tel objet peut être appelé "vecteur scalaire", à condition de bien comprendre de quoi il s'agit. Bien évidemment, si on décide d'inclure le champ dans le système que l'on quantifie, la situation change complètement : alors les sources du champ et le système qui était seul quantifié auparavant forment un et un seul "super-système" quantifié. Toutes les grandeurs dynamiques deviennent alors des observables associées à des opérateurs et pour chacun de ceux-ci, la question de savoir s'ils sont, par exemple, des opérateurs vectoriels, n'est plus triviale.

Il est bien connu que les scalaires et les vecteurs sont les premiers éléments apparaissant lors d'un développement multipolaire. Par exemple, le potentiel électrostatique

[51] Si les états sont en outre caractérisés par d'autres nombres quantiques $\{\alpha\}$ associés aux valeurs propres d'observables commutant avec $\vec{J}^2$, J_z, (18.228) s'écrit :

$$|\alpha jm\rangle' = \sum_{\alpha',\, m'} |\alpha' jm'\rangle\, \langle \alpha' jm'|\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\, \vec{u}.\vec{J}}|\alpha jm\rangle \ . \tag{18.229}$$

Par ailleurs, si $\vec{u} \parallel Oz$, la somme se réduit à un seul terme puisque :

$$\langle \alpha' jm'|\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\, \vec{u}.\vec{J}}|\alpha jm\rangle = \langle \alpha' jm'|\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta J_z}|\alpha jm\rangle = \mathrm{e}^{-\mathrm{i}m\theta}\langle \alpha' jm'|\alpha jm\rangle = \mathrm{e}^{-\mathrm{i}m\theta}\delta_{\alpha\alpha'}\delta_{mm'} \ . \tag{18.230}$$

[52] $\vec{S}$ commutant avec $\vec{L}$, les relations caractéristiques (18.234) ne sauraient être satisfaites avec le couple $(\vec{L}, \vec{S})$.

créé en son extérieur par une certaine distribution $\rho(\vec{r})$ de charge d'extension *finie* commence par un terme où apparaît la charge totale (scalaire), suivi par un autre où figure le moment dipolaire (vecteur). Les termes suivants font ressortir des quantités de type tensoriel impliquant des combinaisons de plus en plus complexes, résultant de la moyenne spatiale avec $\rho(\vec{r})$ de monômes des coordonnées dont le degré croît avec le rang du terme considéré. Ce type de développement survient aussi en Mécanique quantique, le rôle de la densité étant joué d'une façon générale par une quantité du genre $\psi_a^*(\vec{r})\psi_b(\vec{r})$, représentant un certain élément de matrice $\langle\psi_a|\bullet|\psi_b\rangle$; c'est par exemple le cas dans la description semi-classique de l'interaction entre un atome et un champ électromagnétique (voir chapitre 24, sous-section 24.5.1). Pour ce type de question, les arguments de symétrie, tout particulièrement celle liée à la rotation, jouent un rôle essentiel en tant qu'indicateurs des transitions permises ou interdites – éventuellement dans le cadre d'une approximation donnée. Ils débouchent tout naturellement sur la notion de *règle de sélection*, dont un premier exemple sera cité p. 763.

L'intrusion d'objets mathématiques généralisant en ce sens la notion de scalaire et de vecteur justifie la définition d'opérateurs doués de propriétés particulières, plus précisément d'un comportement spécifique par rotation, généralisant précisément celui qui est constaté pour les scalaires ou les vecteurs. Ainsi, c'est le mode de transformation (18.232) qui est généralisé pour poser la définition d'un opérateur tensoriel irréductible[53] (OTI) d'ordre k : les $(2k+1)$ opérateurs $T_q^{(k)}$ (l'indice q variant par pas d'une unité de $-k$ à $+k$) sont par définition les composantes standard d'un opérateur tensoriel irréductible d'ordre k, noté $\mathbf{T}^{(k)}$, s'ils se transforment par rotation suivant la loi ([15], chapitre XIII) :

$$\boxed{(T_q^{(k)})' \stackrel{\text{déf}}{=} \mathcal{R}^{(q)}\,T_q^{(k)}\,\mathcal{R}^{(q)\dagger} = \sum_{q'} T_{q'}^{(k)}\,\mathcal{R}_{q'q}^{(k)}} \tag{18.233}$$

Il n'est pas difficile de montrer ([15], p. 488) que cette définition est équivalente à la satisfaction des relations de commutation suivantes :

$$\boxed{[J_\pm,\,T_q^{(k)}] = \hbar\sqrt{k(k+1)-q(q\pm 1)}\,T_{q\pm 1}^{(k)} \equiv \hbar R_\pm(k,\,q)\,T_{q\pm 1}^{(k)}} \tag{18.234}$$

et :

$$\boxed{[J_z,\,T_q^{(k)}] = \hbar q\,T_q^{(k)}} \tag{18.235}$$

Les opérateurs scalaires sont des OTI d'ordre 0 ; les opérateurs vectoriels sont des OTI d'ordre 1, dont les composantes standard s'expriment à l'aide des composantes cartésiennes selon les formules (17.119).

Ces définitions étant posées, il est alors possible de démontrer l'important théorème de Wigner-Eckart[54], dont l'énoncé est le suivant :

[53]Sur l'origine de la terminologie, voir [15], p. 487.

[54]j et m ont la signification habituelle ; déjà introduit (voir par exemple (18.229)), α est un nombre quantique (ou un ensemble de nombres quantiques) nécessaire pour spécifier complètement l'état quantique, indépendamment des valeurs du moment cinétique. Ce peut être par exemple le nombre quantique n spécifiant la valeur de l'énergie pour le champ central Coulombien.

Dans une représentation $|\alpha jm\rangle$, l'élément de matrice d'une composante standard d'OTI est proportionnel à un coefficient de Clebsch-Gordan[55] *:*

$$\langle \alpha jm|T_q^{(k)}|\alpha' j' m'\rangle = \langle \alpha j\|\mathbf{T}^{(k)}\|\alpha' j'\rangle \, \langle j' k m' q|jm\rangle \ , \qquad (18.237)$$

où $\langle \alpha j\|\mathbf{T}^{(k)}\|\alpha' j'\rangle$ dépend de α, α', j, j' (et de l'OTI !), mais pas de m, m' et q, et s'appelle élément de matrice réduit.

Pour démontrer ce théorème, soit l'élément de matrice des deux relations de définition (18.235) entre $\langle \alpha' j' m'|$ et $|\alpha jm\rangle$; la première donne

$$\langle \alpha jm|[J_z,\, T_q^{(k)}]|\alpha' j' m'\rangle = \hbar q \, \langle \alpha jm|T_q^{(k)}|\alpha' j' m'\rangle \ , \qquad (18.238)$$

Le premier membre est $\hbar(m-m')\langle \alpha jm|T_q^{(k)}|\alpha' j' m'\rangle$, d'où :

$$(m - m' - q)\langle \alpha jm|T_q^{(k)}|\alpha' j' m'\rangle = 0 \ . \qquad (18.239)$$

Prenons maintenant le même élément de matrice pour la relation (18.234) :

$$\langle \alpha jm|[J_\pm,\, T_q^{(k)}]|\alpha' j' m'\rangle = \hbar R_\pm(k,\, q)\langle \alpha jm|T_{q\pm 1}^{(k)}|\alpha' j' m'\rangle \ ; \qquad (18.240)$$

en développant, on obtient :

$$R_\mp(j,\, m)\langle \alpha jm \mp 1|T_q^{(k)}|\alpha' j' m'\rangle = R_\pm(j',\, m')\langle \alpha jm|T_q^{(k)}|\alpha' j' m' \pm 1\rangle + R_\pm(k,\, q)\langle \alpha jm|T_{q\pm 1}^{(k)}|\alpha' j' m'\rangle \ . \qquad (18.241)$$

Récrivons maintenant la relation de récurrence des CG (18.212) en exploitant le fait que tous sont réels :

$$R_\mp(J,\, M)\langle JM \mp 1|j_1 j_2 m_1 m_2\rangle = R_\pm(j_1, m_1)\langle JM|j_1 j_2 m_1 \pm 1 m_2\rangle + R_\pm(j_2 m_2)\langle JM|j_1 j_2 m_1 m_2 \pm 1\rangle \ . \qquad (18.242)$$

La comparaison de (18.241) et (18.242) montre que les CG $\langle JM|j_1 j_2 m_1 m_2\rangle$ et les éléments de matrice $\langle \alpha jm|\, T_q^{(k)}|\alpha' j' m'\rangle$ satisfont exactement les mêmes récurrences linéaires homogènes du premier ordre, avec $j = J$, $m = m'$, $j' = j_1$, $m' = m_1$, $k = j_2$ et $q = m_2$. Il en résulte que $\langle \alpha jm|\, T_q^{(k)}|\alpha' j' m'\rangle$ et $\langle jm|j' k m' q\rangle = \langle j' k m' q|jm\rangle$ sont proportionnels, la constante étant indépendante de m, m' et q ; celle-ci est notée $\langle \alpha j\|\mathbf{T}^{(k)}\|\alpha' j'\rangle$, d'où le théorème de Wigner - Eckart.

La puissance de ce théorème réside dans le fait que l'on connaît les règles suffisantes de nullité de $\langle j' k m' q|jm\rangle$ (voir (18.207)), d'où l'obtention immédiate des règles analogues

[55] Dans la suite, on utilise les notations simplifiées (18.181) ou (18.180) quand c'est possible sans confusion, soit ici :

$$\langle j' k m' q|jmj' k\rangle \equiv \langle j' k m' q|jm\rangle = \langle m' q|jm\rangle \ . \qquad (18.236)$$

Par ailleurs, la définition de l'élément de matrice réduit $\langle \alpha j\|\mathbf{T}^{(k)}\|\alpha' j'\rangle$ diffère d'un auteur à l'autre ; par exemple, Messiah ([15] p. 489) introduit explicitement un facteur $\frac{1}{\sqrt{2j+1}}$ dans l'énoncé du théorème (18.237).

pour l'élément de matrice $\langle\alpha jm|T_q^{(k)}|\alpha' j'm'\rangle$. Noter aussi que toute la dépendance de ce dernier vis-à-vis des nombres magnétiques est exclusivement contenue dans le coefficient CG à droite.

Donnons un exemple de l'importance et de l'utilité de ce théorème en l'appliquant dans quelques cas simples. On sait que le produit scalaire $\vec{V}.\vec{W}$ de deux opérateurs vectoriels est un opérateur scalaire (voir exercice 17.8.1), caractérisé par $k = q = 0$; le théorème donne alors :

$$\langle\alpha jm|\vec{V}.\vec{W}|\alpha' j'm'\rangle = \langle\alpha\, j\|\vec{V}.\vec{W}\|\alpha'\, j'\rangle\,\langle j'0m'0|jm\rangle\ . \tag{18.243}$$

Le CG de droite est nul si $j \neq j'$ et/ou $m \neq m'$, et vaut visiblement 1 quand il n'est pas nul ; on trouve ainsi :

$$\langle\alpha jm|\vec{V}.\vec{W}|\alpha' j'm'\rangle = \langle\alpha\, j\|\vec{V}.\vec{W}\|\alpha'\, j'\rangle\,\delta_{jj'}\delta_{mm'} \tag{18.244}$$

qui reproduit les deux évidences : $\langle\alpha jm|\vec{V}.\vec{W}|\alpha' j'm'\rangle$ est nul entre deux états propres de $(\vec{J}^2,\ J_z)$ distincts, et est indépendant de m dans le cas contraire. Prenant $\vec{V} = \vec{W} = \vec{J}$, on obtient aussi :

$$\langle\alpha\, j\|\vec{J}^2\|\alpha'\, j\rangle\ = \langle\alpha jm|\vec{J}^2|\alpha' jm\rangle = \delta_{\alpha\alpha'}\hbar^2 j(j+1)\ . \tag{18.245}$$

Appliquons maintenant le théorème aux deux opérateurs scalaires $\vec{V}.\vec{J}$ et $\vec{J}^2$:

$$\langle\alpha jm|\vec{V}.\vec{J}|\alpha' jm\rangle = \langle\alpha\, j\|\vec{V}.\vec{J}\|\alpha'\, j\rangle\,\langle j0m0|jm\rangle\ , \tag{18.246}$$

$$\langle\alpha jm|\vec{J}^2|\alpha' jm\rangle = \langle\alpha\, j\|\vec{J}^2\|\alpha'\, j\rangle\,\langle j0m0|jm\rangle\ ; \tag{18.247}$$

le rapport membre à membre donne :

$$\langle\alpha jm|\vec{V}.\vec{J}|\alpha' jm\rangle = \frac{\langle\alpha\, j\|\vec{V}.\vec{J}\|\alpha'\, j\rangle}{\langle\alpha\, j\|\vec{J}^2\|\alpha'\, j\rangle}\langle\alpha jm|\vec{J}^2|\alpha' jm\rangle = \langle\alpha\, j\|\vec{V}.\vec{J}\|\alpha'\, j\rangle\ , \tag{18.248}$$

la dernière égalité résultant de (18.245). D'un autre côté, l'élément de matrice de gauche est (voir (17.121)) :

$$\begin{aligned}\langle\alpha jm|\vec{V}.\vec{J}|\alpha' jm\rangle = \hbar\big[m\langle\alpha jm|V_0^{(1)}|\alpha jm\rangle - R_-(j,\, m)\langle\alpha jm|V_1^{(1)}|\alpha jm-1\rangle \\ -R_+(j,\, m)\langle\alpha jm|V_{-1}^{(1)}|\alpha jm+1\rangle\big]\ ,\end{aligned} \tag{18.249}$$

mais, toujours d'après le théorème de Wigner - Eckart, chacun des éléments de matrice situés dans les crochets est proportionnel à $\langle\alpha j\|\vec{V}\|\alpha' j\rangle$. Au total, $\langle\alpha jm|\vec{V}.\vec{J}|\alpha' jm\rangle$ est de la forme $C\langle\alpha j\|\vec{V}\|\alpha' j\rangle$, où C est une constante indépendante de $\vec{V}$; faisant $\vec{V} = \vec{J}$, on obtient $\langle\alpha jm|\vec{J}^2|\alpha' jm\rangle = C\langle\alpha j\|\vec{J}\|\alpha' j\rangle$, soit $C = \delta_{\alpha\alpha'}\hbar^2 j(j+1)/\langle\alpha j\|\vec{J}\|\alpha' j\rangle$. L'égalité (18.248), lue à l'envers prend alors la forme :

$$\langle\alpha\, j\|\vec{V}.\vec{J}\|\alpha'\, j\rangle = C\langle\alpha j\|\vec{V}\|\alpha' j\rangle = \frac{\delta_{\alpha\alpha'}\hbar^2 j(j+1)}{\langle\alpha j\|\vec{J}\|\alpha' j\rangle}\,\langle\alpha j\|\vec{V}\|\alpha' j\rangle\ ; \tag{18.250}$$

avec $\alpha = \alpha'$, il vient :

$$\boxed{\langle \alpha\, j \| \vec{V} \| \alpha\, j \rangle = \frac{\langle \alpha jm | \vec{V}.\vec{J}\,| \alpha jm \rangle}{\hbar^2 j(j+1)}\, \langle \alpha\, j \| \vec{J}\,\| \alpha\, j \rangle} \tag{18.251}$$

Cette proportionnalité entre tout opérateur vectoriel et le moment cinétique est le fondement théorique du modèle vectoriel de l'atome, qui permet de raisonner physiquement en terme de précession des moments cinétiques. L'égalité (18.251) est suffisamment importante pour mériter une démonstration directe spécifique ; elle est proposée dans le problème 18.7.6.

Considérons maintenant le cas des éléments de matrice d'opérateurs vectoriels, à commencer par le plus fondamental d'entre eux, $\vec{J}$: $\vec{J} \equiv \mathbf{T}^{(1)}$. D'après (17.119), ses composantes standard sont :

$$J_{+1}^{(1)} = -\frac{1}{\sqrt{2}}\,(J_x + \mathrm{i}\, J_y) \ , \qquad J_{-1}^{(1)} = +\frac{1}{\sqrt{2}}\,(J_x - \mathrm{i}\, J_y) \ , \qquad J_0^{(1)} = J_z \ . \tag{18.252}$$

Par le théorème de Wigner-Eckart, (18.237), on a :

$$\langle \alpha jm | J_z | \alpha' j' m' \rangle = \langle \alpha j \| \vec{J}\, \| \alpha' j' \rangle \, \langle j' 1 m' 0 | jm \rangle \ , \tag{18.253}$$

où $\vec{J}$ au centre de l'élément de matrice réduit est donc l'équivalent de $\mathbf{T}^{(1)}$. Si $j' \neq j,\ j \pm 1$, et/ou $m \neq m'$, le CG est nul ; on peut donc écrire, après simplification par $\delta_{mm'}$:

$$\begin{aligned} m\delta_{\alpha\alpha'}\delta_{jj'} = \langle \alpha j \| \vec{J}\, \| \alpha' j' \rangle \, [& \langle j1m0|jm\rangle \delta_{j'j} + \\ & \langle j-11m0|jm\rangle \delta_{j'j-1} + \langle j+11m0|jm\rangle \delta_{j'j+1}] \ , \end{aligned} \tag{18.254}$$

Si $j' = j \pm 1$, on obtient $\langle \alpha j \| \vec{J}\, \| \alpha' j \pm 1 \rangle = 0$; si $j' = j$, il reste :

$$m\hbar\delta_{\alpha\alpha'} = \langle \alpha j \| \vec{J}\, \| \alpha' j \rangle \, \langle j1m0|jm\rangle \ . \tag{18.255}$$

La dépendance en m n'est qu'apparente : si $m = 0$, le CG à droite est nul ; par ailleurs, en vertu de (18.219), on a $\langle j1-m0|j-m\rangle = -\langle j1m0|jm\rangle$: passer de $+m$ à $-m$ ne fait que changer le signe des deux membres. Plus généralement (voir exercice 18.7.5) :

$$\langle j1m0|jm\rangle = \frac{m}{\sqrt{j(j+1}} \ . \tag{18.256}$$

de sorte que finalement :

$$\boxed{\langle \alpha j \| \vec{J}\, \| \alpha' j' \rangle = \delta_{\alpha\alpha'}\delta_{jj'}\hbar\sqrt{j(j+1} \equiv \delta_{\alpha\alpha'}\delta_{jj'}\, \|\vec{J}\,\|} \tag{18.257}$$

Cette égalité est vraie pour les trois composantes standard $J_q^{(1)}$ de $\vec{J}$ définies en (18.252).

Ces résultats permettent de déterminer complètement les éléments de matrice d'un opérateur vectoriel *quelconque* $\vec{V}$, dans un sous-espace $\{|\alpha jm\rangle\}_m$ donné, en fonction de

ceux de $\vec{J}$, qui sont connus. Écrivons (18.237) pour $\vec{V}$ et $\vec{J}$, et faisons le rapport membre à membre ; le coefficient CG s'élimine, et il reste :

$$\langle\alpha jm|V_q^{(1)}|\alpha jm'\rangle = \frac{\langle\alpha\, j\|\vec{V}\|\alpha\, j\rangle}{\langle\alpha\, j\|\vec{J}\,\|\alpha\, j\rangle}\,\langle\alpha jm|J_q^{(1)}|\alpha jm'\rangle = \frac{\langle\alpha\, j\|\vec{V}\|\alpha\, j\rangle}{\hbar\sqrt{j(j+1)}}\,\langle\alpha jm|J_q^{(1)}|\alpha jm'\rangle\ . \tag{18.258}$$

Compte tenu de (18.251), ceci se récrit :

$$\langle\alpha jm|V_q^{(1)}|\alpha jm'\rangle = \frac{\langle\alpha\, jm|\vec{V}.\vec{J}|\alpha\, jm\rangle}{\hbar^2 j(j+1)}\,\langle\alpha jm|J_q^{(1)}|\alpha jm'\rangle\ . \tag{18.259}$$

Comme déjà relevé, le numérateur est indépendant de m (voir (18.244)) et représente la valeur moyenne de l'opérateur scalaire $\vec{V}.\vec{J}$ dans le sous-espace $\{\alpha jm\rangle\}_m$, ce que l'on peut noter $\langle\vec{V}.\vec{J}\rangle_{\alpha j}$. Alors, (18.259) prend la forme classique :

$$\boxed{\langle\alpha jm|\vec{V}|\alpha jm'\rangle = \frac{\langle\vec{V}.\vec{J}\rangle_{\alpha j}}{\hbar^2 j(j+1)}\,\langle\alpha jm|\vec{J}\,|\alpha jm'\rangle} \tag{18.260}$$

et porte, pour des raisons géométriques et algébriques évidentes, le nom de *Théorème de projection.* Géométriques car on y voit comme un vecteur unitaire $\frac{\vec{J}}{\|\vec{J}\|}$ dont on prend le produit scalaire avec $\vec{V}$ pour définir une projection de $\vec{V}$ le long de $\vec{J}$; algébriques aussi, car cette relation ne tient que dans un sous-espace $\{|\alpha jm\rangle\}_m$ donné, et concerne donc fondamentalement l'opérateur projeté $P_{\alpha j}\vec{V}P_{\alpha j}$ où $P_{\alpha j}$ est le projecteur $\sum_m |\alpha jm\rangle\langle\alpha jm|$.

À titre d'illustration physique, considérons pour finir le moment magnétique d'un atome, qui résulte tout à la fois du moment cinétique orbital et du moment cinétique intrinsèque (spin) ; on peut donc l'écrire en combinaison linéaire des opérateurs vectoriels correspondants. Il se trouve que, outre le facteur gyromagnétique habituel, un facteur numérique[56] g_e doit être introduit en ce qui concerne le spin (c'est l'expérience qui l'impose, voir chapitre 20), obligeant à écrire comme suit le moment magnétique :

$$\vec{\mu} = \frac{e}{2m}\,(\vec{L} + g_\mathrm{e}\vec{S})\ . \tag{18.261}$$

$\vec{L}$ et $\vec{S}$ étant des opérateurs vectoriels, $\vec{\mu}$ en est un aussi. D'après (18.260), on a :

$$\langle\alpha jm|\vec{\mu}|\alpha jm'\rangle = \frac{e}{2m}\,\frac{\langle(\vec{L}+g_\mathrm{e}\vec{S}).(\vec{L}+\vec{S}\,)\,\rangle_{\alpha j}}{\hbar^2 j(j+1)}\,\langle\alpha jm|\vec{J}\,|\alpha jm'\rangle \tag{18.262}$$

Le produit scalaire au numérateur s'écrit $\vec{L}^{\,2} + g_\mathrm{e}\vec{S}^{\,2} + \frac{1}{2}(g_\mathrm{e}+1)(\vec{J}^{\,2} - \vec{L}^{\,2} - \vec{S}^{\,2})$, et vaut $\frac{1}{2}[(1+g_\mathrm{e})\vec{J}^{\,2} + (1-g_\mathrm{e})(\vec{L}^{\,2} - \vec{S}^{\,2})]$; prenant $g_\mathrm{e} = 2$, et avec[57] $|jm\rangle \equiv |jmls\rangle$, on obtient :

$$\boxed{\langle\alpha jm|\vec{\mu}|\alpha jm'\rangle = g_{lsj}\,\frac{e}{2m}\,\langle\alpha jm|\vec{J}\,|\alpha jm'\rangle} \tag{18.263}$$

[56] $g_\mathrm{e} = 2,002\,319\,304\,362\,2...$; en pratique, on prend le plus souvent $g_\mathrm{e} = 2$. Ces points seront discutés dans le chapitre 20.

[57] C'est la bonne base pour le couplage dit $\vec{L}.\vec{S}$, ou *de Russel - Saunders*, décrivant convenablement la structure fine pour les éléments légers (voir chapitre 27).

où g_{lsj} est le *facteur de Landé* :

$$\boxed{g_{lsj} \stackrel{\text{déf}}{=} 1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)}} \qquad (18.264)$$

En l'absence de spin, ou si $s = 0$ (état singulet), j est égal à l, d'où $g_{lsj} = 1$: en pareil cas, on retrouve pour les éléments de matrice la proportionnalité élémentaire avec le facteur gyromagnétique classique $\gamma \stackrel{\text{déf}}{=} \frac{e}{2m}$.

18.7 Exercices et problèmes

18.7.1 Le vecteur $\vec{L}$ en coordonnées sphériques

1. Établir les expressions (18.51) et (18.50) des composantes du moment cinétique orbital $\vec{L} = \vec{r} \times \vec{p}$ en coordonnées sphériques (représentation-q).
2. Vérifier qu'elles satisfont les relations de commutation caractéristiques (18.2).
3. Montrer que L_z est hermitique dans l'espace des fonctions 2π-périodiques.
4. Trouver les commutateurs du moment cinétique orbital avec la coordonnée $\vec{r}$ et l'impulsion $\vec{p}$. Montrer explicitement que $\vec{r}^{\,2}$ et $\vec{p}^{\,2}$ sont des opérateurs scalaires.
5. En repartant de l'expression de $L_\pm$ en coordonnées sphériques, et en raisonnant avec des fonctions 2π-périodiques, montrer que $L_-^\dagger = L_+$.
6. Écrire les expressions des L_{p_u} en représentation-p.

18.7.2 Quantification d'une variable angulaire

Il s'agit d'examiner les subtilités [28] que posent les variables angulaires vis-à-vis des relations de commutation canoniques. Pour fixer les idées, on considère l'angle azimutal ϕ et son moment conjugué $p_\phi \equiv L_z$, égal à l'opérateur différentiel $-\mathrm{i}\hbar\partial_\phi$ en représentation-q. Dans la première partie, on commet délibérément une erreur de fond, et on met en évidence une contradiction ; dans la deuxième partie, on rectifie le tir, et on complète le contenu de la Remarque p. 728.

1. Dans cette partie, on commet l'erreur consistant à poser :

$$[\phi,\, L_z] = -\mathrm{i}\hbar\,\mathbf{1} \qquad \textbf{(faux !)} \ , \qquad (18.265)$$

raisonnant trop vite en représentant l'opérateur associé à l'angle ϕ par la simple multiplication[58] par ϕ :

$$\phi \leftrightarrow \hat{\phi} \rightarrow \times \phi \qquad \textbf{(faux !)} \tag{18.266}$$

en prenant l'angle ϕ dans l'intervalle $[0,\ 2\pi]$.

(a) En admettant l'association (18.266), montrer que pour toute fonction $f(\phi)$ on a :

$$\frac{\partial}{\partial\phi}\hat{\phi}f(\phi) - \hat{\phi}\frac{\partial}{\partial\phi}f(\phi) = f(\phi) \ . \tag{18.267}$$

et en déduire alors l'égalité (18.265).

(b) Selon la théorie des séries de Fourier (et le théorème de Dirichlet), l'ensemble des fonctions $e_n(\phi) \stackrel{\text{déf}}{=} \mathrm{e}^{in\phi}$, $n \in \mathbb{Z}$, constitue une base complète pour l'espace des fonctions 2π-périodiques à variation bornée. Montrer que ces fonctions sont deux à deux orthogonales, et les normaliser à l'unité, $\langle e_n|e_m\rangle = \delta_{nm}$.

(c) Calculer les éléments de matrice $\langle e_n|\hat{\phi}|e_m\rangle$ et $\langle e_n|L_z|e_m\rangle$.

(d) En effectuant la multiplication des matrices, trouver l'élément de matrice de $\langle e_n|[\hat{\phi},\ L_z]|e_m\rangle$. Par comparaison avec (18.265), mettre en évidence la contradiction.

2. Dans cette partie, on définit convenablement l'opérateur $\hat{\phi}$, comme indiqué dans la Remarque p. 728.

(a) Tracer le graphe de la fonction périodique $\phi_{\text{per}}(\phi)$ définie en (18.55). Expliquer pourquoi sa dérivée contient un peigne de Dirac.

(b) Calculer la série de Fourier de $\phi_{\text{per}}(\phi)$, et retrouver analytiquement le peigne de Dirac.

(c) En déduire l'égalité (18.57).

(d) Combien vaut l'élément de matrice $\langle e_n|\hat{\phi}|e_m\rangle$?

(e) En prenant l'élément de matrice membre à membre de l' égalité (18.57), montrer que :

$$\langle e_n|[\hat{\phi},\ L_z]|e_m\rangle = \mathrm{i}\hbar(\delta_{nm} - 1) \ , \tag{18.268}$$

et commenter en comparant avec le résultat 1d.

(f) En utilisant le développement de Fourier, retrouver pour $\langle e_n|\hat{\phi}|e_m\rangle$ le résultat (correct) obtenu en 1c.

(g) Si on définit les écarts quadratiques $\Delta\phi^2$ et ΔL_z^2 comme d'habitude, a-t-on la relation[59] $\Delta\phi\Delta L_z \geq \frac{\hbar}{2}$?

(Pour en savoir plus sur ce sujet, la lecture des articles de Judge et Lewis [27] et de Lévy-Leblond [28] est recommandée).

[58] Comparer (18.266) à (18.55).
[59] Voir Tome I, éq. (10.97).

18.7.3 Quelques résultats à propos d'un moment cinétique

Au besoin, on désigne par $|j\,m\rangle$ un état propre du couple $(\vec{J}^2,\ J_z)$.

1. Montrer que si un opérateur commute avec deux composantes de $\vec{J}$, il commute aussi avec la troisième[60].

2. Démontrer l'égalité :

$$|j\,m\rangle = \hbar^{m-j}\sqrt{\frac{(j+m)!}{(2j)!(j-m)!}}\ J_-^{j-m}\,|j\,j\rangle\ , \tag{18.269}$$

   En déduire l'élément de matrice $\langle j\,j|J_+^{j-m}|j\,m\rangle$.

3. Soit $\vec{V}$ un opérateur vectoriel ; $\vec{n}$ et $\vec{u}$ désignant deux vecteurs "scalaires", montrer que :

$$[\vec{u}.\vec{J},\ \vec{n}.\vec{V}\,] = \mathrm{i}\hbar\,(\vec{u}\times\vec{n}).\vec{V}\ . \tag{18.270}$$

   En déduire les relations caractéristiques du moment cinétique sous la forme :

$$[\vec{a}.\vec{J},\ \vec{b}.\vec{J}] = \mathrm{i}\hbar\,\vec{J}.(\vec{a}\times\vec{b})\ . \tag{18.271}$$

4. Simplifier l'expression $\mathrm{e}^{\mathrm{i}\pi J_x/(2\hbar)}\,\mathrm{e}^{-\mathrm{i}\beta J_z/\hbar}\,\mathrm{e}^{-\mathrm{i}\pi J_x/(2\hbar)}$.

5. Démontrer l'égalité (18.143).

6. Pour $j=1$, trouver les matrices représentant J_x, J_y, J_z, $J_\pm$, J_x^2, et $\vec{J}^2$.

7. Établir les expressions (18.44) des incertitudes ΔJ_x et ΔJ_y dans un état $|j\,m\rangle$.

8. Pour j quelconque, trouver les éléments de matrice $\langle j',m'|J_u|j\,m\rangle$, $u=x,y$. En déduire, pour un état $|j\,m\rangle$, la valeur moyenne et l'écart quadratique de $\vec{J}.\vec{n}$, $\vec{n}$ étant le vecteur unitaire de la direction $(\theta,\ \phi)$.

9. On note simplement $|m\rangle$ $(m=0,\pm1)$ les trois états propres d'un moment cinétique $j=1$. Trouver les valeurs moyennes de J_x, J_y, J_z, J_x^2, J_y^2 et J_z^2 dans l'état normalisé $|\psi\rangle = a\,|+1\rangle + b\,|0\rangle + c\,|-1\rangle$ (a, b et c complexes quelconques).

18.7.4 Moment cinétique $j=\frac{1}{2}$

Un moment cinétique $j=1/2$ peut s'écrire $\vec{J}=(\hbar/2)\,\vec{\sigma}$ où $\vec{\sigma}$ désigne les trois matrices de Pauli : $\vec{\sigma}\equiv(\sigma_x,\ \sigma_y,\ \sigma_z)$.

[60] Ainsi, pour établir qu'un opérateur est scalaire, il suffit de démontrer qu'il commute avec deux des composantes du moment cinétique.

1. Calculer les commutateurs $[\sigma_u,\, \sigma_v]$. Combien vaut σ_u^2 ? Et $\vec{\sigma}^{\,2}$?
2. Écrire les matrices σ_u.
3. Calculer les anticommutateurs $\{\sigma_u,\, \sigma_v\}$; en déduire les produits $\sigma_u\sigma_v$.
4. Combien valent les traces des σ_u ?
5. Toute matrice 2×2, M, peut s'écrire sous la forme $M = m_0\, \mathbf{1} + \vec{m}.\vec{\sigma}$. Exprimer m_0 et $\vec{m}$ sous la forme d'une trace.
6. Démontrer la formule de Moivre quantique :

$$\mathrm{e}^{\mathrm{i}\phi\, \vec{n}.\vec{\sigma}} = \cos\phi\, \mathbf{1}_2 + \mathrm{i}\, \vec{n}.\vec{\sigma}.\sin\phi \tag{18.272}$$

   où ϕ est un complexe quelconque et $\vec{n}$ un vecteur unitaire scalaire.

18.7.5 Calcul de $\langle j1j0|jj\rangle$ et démonstration de $\langle j100|j0\rangle = 0$

1. Calculer le coefficient de Clebsch - Gordan $\langle j1j0|jj\rangle$; plus généralement, montrer que $\langle j1m0|jm\rangle$ est donné par (18.256).
2. Montrer que $\langle j100|j0\rangle = 0$.

18.7.6 Le théorème de Wigner - Eckart pour les opérateurs vectoriels

Le but de cet exercice est de démontrer directement le théorème de Wigner - Eckart pour les opérateurs vectoriels. Dans la suite, on considère exclusivement le sous-espace propre $\mathcal{E}_{\tau j}$ d'un ECOC $(T_i,\, \vec{J}^2,\, J_z)$ contenant le moment cinétique, engendré par les $2j+1$ vecteurs $\{|\tau, j, m\rangle\}$, $-j \leq m \leq +j$. τ désigne l'ensemble des nombres quantiques associés aux observables T_i.

1. Soit $\vec{V}$ un opérateur vectoriel, satisfaisant par définition les relations (17.117) avec $\vec{L} \equiv \vec{J}$, pour la généralité des écritures.
   (a) Montrer que $\vec{V}^{\,2}$ est un opérateur scalaire.
   (b) On pose $V_\pm = V_x \pm \mathrm{i}\, V_y$. Calculer $[J_u,\, V_\pm]$ $(u = x, y, z)$, $[J_\pm, V_\pm]$, $[J_-, V_+]$ et $[J_+, V_-]$.
   (c) Montrer que si l'on applique $V_\pm$ et V_z à un vecteur $|\tau, j, m\rangle$, on obtient un vecteur propre de J_z ; quelles sont les valeurs propres associées ?
   (d) En déduire que $\langle \tau, j, m'|V_\pm|\tau\, j\, m\rangle$ est nul si $m - m' \neq \pm 1$.

2. (a) Calculer l'élément de matrice $\langle \tau, j, m+2|[J_+, V_+]|\tau, j, m\rangle$ et montrer que :

$$\langle \tau\, j\, m+2|J_+|\tau\, j\, m+1\rangle\langle \tau\, j\, m+1|V_+|\tau\, j\, m\rangle = \langle \tau,\, j,\, m+1|J_+|\tau\, j\, m\rangle\langle \tau\, j\, m+2|V_+|\tau\, j\, m+1\rangle \;. \quad (18.273)$$

(b) En déduire que, dans le sous-espace $\mathcal{E}_{\tau j}$, les éléments de matrice de V_+ sont proportionnels à ceux de J_+ :

$$\langle \tau\, j\, m|V_+|\tau\, j\, m'\rangle = \alpha_+(\tau,\, j)\langle \tau\, j\, m|J_+|\tau\, j\, m'\rangle \;. \quad (18.274)$$

(c) En procédant de façon analogue, montrer que :

$$\langle \tau\, j\, m|V_-|\tau,\, j,\, m'\rangle = \alpha_-(\tau,\, j)\,\langle \tau,\, j,\, m|J_-|\tau,\, j,\, m'\rangle \;. \quad (18.275)$$

(d) A l'aide des commutateurs $[J_+,\, V_-]$ et $[J_-, V_+]$, montrer que[61] :

$$\langle \tau,\, j,\, m|V_z|\tau\, j\, m\rangle = m\hbar\alpha_+(\tau,\, j) \;\;,\;\; \langle \tau\, j\, m|V_z|\tau\, j\, m\rangle = m\hbar\alpha_-(\tau,\, j) \;. \quad (18.276)$$

(e) Les résultats précédents montrent que, à l'intérieur de $\mathcal{E}_{\tau j}$, $\vec{V}$ et $\vec{J}$ ont des matrices proportionnelles. En notant $P_{\tau j}$ le projecteur dans $\mathcal{E}_{\tau j}$, on peut donc écrire :

$$P_{\tau j}\, \vec{V}\, P_{\tau j} = \alpha(\tau,\, j)\, P_{\tau j}\, \vec{J}\, P_{\tau j} \;. \quad (18.277)$$

Calculer la valeur moyenne du produit scalaire $\vec{V}.\vec{J}$ dans un état $|\psi\rangle$ quelconque appartenant à $\mathcal{E}_{\tau j}$; constatant qu'elle est indépendante de $|\psi\rangle$, on la note $\langle \vec{V}.\vec{J}\rangle_{\tau j}$. En déduire que :

$$P_{\tau j}\, \vec{V}\, P_{\tau j} = \frac{\langle \vec{V}.\vec{J}\rangle_{\tau j}}{\hbar^2\, j(j+1)}\, P_{\tau j}\, \vec{J}\, P_{\tau j} \quad (18.278)$$

et interpréter cette relation[62].

18.7.7 Addition de deux moments cinétiques

On considère deux moments cinétiques $j_1 = 1$ et $j_2 = 1/2$. Les états propres communs à $(\vec{J}_1^2,\, J_{1z},\, \vec{J}_2^2,\, J_{2z})$ sont notés $|j_1\, j_2\, m_1\, m_2\rangle$.

1. Quelle est la dimension (multiplicité) de l'espace des états ? Quelles sont les valeurs possibles j du moment cinétique total $\vec{J} = \vec{J}_1 + \vec{J}_2$? Quelle est la dimension de l'espace des états de $\vec{J}$?
2. Déterminer quelques composantes $|jmj_1j_2\rangle$ des états propres de $(\vec{J}^2,\, J_z)$ sous la forme de combinaisons linéaires des états propres $|j_1j_2m_1m_2\rangle$.
3. Trouver toutes les composantes dans le cas particulier $j_1 = j_2 = 1$.

[61] Dans la suite, on note $\alpha_+ = \alpha_- = \alpha$.

[62] Elle constitue la justification théorique du modèle vectoriel de l'atome.

18.7.8 Moment cinétique total de N spins $\frac{1}{2}$. Diagramme de branchement.

On considère N moments cinétiques $j_k = 1/2$ $(1 \leq k \leq N)$; m note la valeur propre associée à $J_z = \sum_{k=1}^{N} J_{k\,z}$.

1. Combien y a-t-il de façons d'avoir $m = N/2$? En déduire le nombre de multiplets $j = N/2$. De même, combien y a-t-il de multiplets $j = (N/2) - 1$?
2. En déduire le nombre de multiplets ayant $j = (N/2) - p$ $(p \leq N/2)$.
3. Pour $N = 4$, combien existe-t-il de quintuplets, de triplets et de singulets ?

18.7.9 Oscillateur harmonique à deux dimensions

Soit un oscillateur harmonique isotrope (masse μ, pulsation ω) à deux dimensions ; l'axe Oz est l'axe perpendiculaire au plan d'oscillation. On définit les deux opérateurs a_u $(u = x, y)$:

$$a_u = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\,u + \mathrm{i}\,\frac{p_u}{\sqrt{\mu\hbar\omega}}\right) \tag{18.279}$$

ainsi que leurs adjoints $a_u^\dagger$. Le Hamiltonien est noté H, les opérateurs nombres de particules sont $N_u \stackrel{\text{déf}}{=} a_u a_u^\dagger$.

1. (a) Montrer que les seuls commutateurs non-nuls sont les $[a_u,\, a_u^\dagger]$ et qu'ils valent $\mathbf{1}$.
   (b) Exprimer H et L_z à l'aide des opérateurs a_u et $a_u^\dagger$. Expliquer pourquoi L_z commute avec H et le vérifier explicitement.
   (c) Soit $|n_x, n_y\rangle$ un état propre de H construit comme le produit tensoriel des états propres de N_x et de N_y ; on pose $n_x + n_y = n$: quelle est la dégénérescence des états de ce type ?
2. On introduit les opérateurs droit et gauche définis comme :

$$a_{\mathrm{d}} \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}(a_x - \mathrm{i}a_y) \qquad a_{\mathrm{g}} \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}(a_x + \mathrm{i}a_y) \tag{18.280}$$

   ainsi que leurs adjoints $a_{\mathrm{d}}^\dagger$ et $a_{\mathrm{g}}^\dagger$.

   (a) Calculer les commutateurs $[a_{\mathrm{c}},\, a_{\mathrm{c}}^\dagger]$ (c = d, g) ; que valent les autres ?
   (b) Exprimer H et L_z en fonction des $N_{\mathrm{c}} = a_{\mathrm{c}}^\dagger a_{\mathrm{c}}$.
   (c) On note désormais $|n, m \stackrel{\text{déf}}{=} n_{\mathrm{d}} - n_{\mathrm{g}}\rangle$ un état propre commun à N_{d} et N_{g}. Montrer que H et L_z forment un ECOC.

(d) Sachant que le fondamental $|n = 0, m = 0\rangle$ est la gaussienne :

$$\psi_{00}(x, y) = C\,\mathrm{e}^{-\alpha(x^2+y^2)/2} \ , \quad C = \left(\frac{\alpha}{\pi}\right)^{1/2} \ , \quad \alpha = \frac{\mu\omega}{\hbar} \ , \tag{18.281}$$

trouver les états $\psi_{1\pm1}$ (les exprimer en fonction des coordonnées cylindriques ρ et ϕ).

(e) Trouver les composantes cylindriques du vecteur courant $\vec{j}$, j_ρ et j_ϕ, sachant que les composantes du gradient dans le plan xOy sont données par :

$$\vec{\nabla} = \left(\frac{\partial}{\partial\rho}, \frac{1}{\rho}\frac{\partial}{\partial\phi}\right) \ . \tag{18.282}$$

(f) L'oscillateur étant à $t = 0$ dans l'état $|\Psi(t = 0)\rangle = \cos\theta|0, 0\rangle + \sin\theta|1, 1\rangle$, calculer les valeurs moyennes des coordonnées x et y à l'instant t.

18.7.10 Matrices de Pauli et vecteur polarisation

On considère les opérateurs de spin $\vec{S}^2$ et S_z et on note $|s, m\rangle$ leurs vecteurs propres communs, tels que :

$$\vec{S}^2|s, m\rangle = s(s+1)\hbar^2|s, m\rangle \ , \qquad S_z|s, m\rangle = m\hbar|s, m\rangle \ . \tag{18.283}$$

Avec $s = \frac{1}{2}$, on pose $|\frac{1}{2}, \pm\frac{1}{2}\rangle = |\pm\rangle$. Les σ_u désignent ci-dessous les trois matrices de Pauli.

1. (a) Que valent les traces $\mathrm{Tr}\sigma_u$ et $\mathrm{Tr}(\sigma_u\sigma_v)$ (distinguer les deux cas $u \neq v$ et $u = v$) ?

   (b) Toute matrice A représentant un opérateur dans la base $|\pm\rangle$ peut-être exprimée d'une manière unique sous la forme d'une combinaison linéaire des σ_u et de la matrice identité $\mathbf{1}$:

   $$A = \frac{1}{2}(\alpha_0\mathbf{1}_2 + \vec{\alpha}.\vec{\sigma}) \tag{18.284}$$

   Montrer que $\alpha_0 = \mathrm{Tr}\ A$ et que $\vec{\alpha} = \mathrm{Tr}(A\,\vec{\sigma})$.

2. Soit une particule de spin 1/2, de Hamiltonien H, dont le vecteur d'état *normé* $|\Psi(t)\rangle$ est développé sur la base $|+\rangle$, $|-\rangle$:

   $$|\Psi\rangle \ = \ a_+|+\rangle + a_-|-\rangle \ , \tag{18.285}$$

   où les $a_\pm$ sont des nombres complexes.

   (a) Soit l'opérateur $P = |\Psi\rangle\langle\Psi|$. À quoi est égal P^2 ? Donner la matrice de P dans la base $|+\rangle$, $|-\rangle$ en fonction des coefficients $a_\pm$.

(b) On veut trouver l'expression de P sous la forme de l'équation (18.284) soit :

$$P = \frac{1}{2}(M_0 \mathbf{1}_2 + \vec{M}.\vec{\sigma}) \ . \tag{18.286}$$

$\vec{M}$ est appelé vecteur polarisation. Montrer que $\mathrm{Tr}P = 1$ et en déduire M_0.

(c) Compte tenu de la relation entre P^2 et P, et de l'expression (18.286), montrer que $\|\vec{M}\| = 1$.

(d) Exprimer $\vec{M}$ en fonction de $\langle\vec{\sigma}\rangle \overset{\text{déf}}{=} \langle\Psi|\vec{\sigma}|\Psi\rangle$.

(e) Calculer les composantes du vecteur $\vec{M}$ en fonction des coefficients a_+ et a_-.

(f) Calculer le vecteur polarisation $\vec{M}$ dans les deux cas particuliers :

$$a_+ = 1 \ , \quad a_- = 0 \qquad \text{et} \qquad a_+ = 0 \ , \quad a_- = 1 \ . \tag{18.287}$$

(g) Compte tenu de l'expression de $\vec{M}$ en fonction de $\langle\vec{\sigma}\rangle$, montrer que l'équation décrivant l'évolution de $\vec{M}$ au cours du temps s'écrit :

$$\frac{\mathrm{d}\vec{M}}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar}\langle\Psi|[H, \vec{\sigma}]|\Psi\rangle \ . \tag{18.288}$$

(h) Le Hamiltonien H, supposé de trace nulle, est décomposé de la même façon et s'écrit $H = \frac{\hbar}{2}\vec{\omega}.\vec{\sigma}$, où $\vec{\omega}$ est supposé connu. Montrer que $\vec{M}$ obéit à une équation de type gyroscopique.

18.7.11 Dynamique d'un système à deux niveaux. Oscillation de Rabi

Soit un système à deux niveaux $|1\rangle$ et $|2\rangle$ séparés en énergie de $\hbar\omega_0$ (voir fig. 18.1) ; le Hamiltonien est :

$$H_0 = \frac{\hbar\omega_0}{2}\left(-|1\rangle\langle 1| + |2\rangle\langle 2|\right) \ . \tag{18.289}$$

Le système étant dans l'état fondamental, on applique à partir de $t = 0$ une perturbation $V(t)$ monochromatique donnée par :

$$V(t) = v(\mathrm{e}^{\mathrm{i}\omega t}\,|1\rangle\langle 2| + \mathrm{e}^{-\mathrm{i}\omega t}\,|2\rangle\langle 1|) \qquad (v \equiv \hbar\Omega \in \mathbb{R}) \ . \tag{18.290}$$

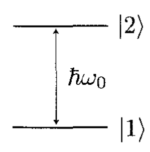


Figure 18.1: Système à deux niveaux

1. Montrer que $H(t) \stackrel{\text{déf}}{=} H_0 + V(t)$ peut se mettre sous la forme $H(t) = \mathcal{R}(t)\tilde{H}\mathcal{R}^\dagger(t)$, où $\tilde{H}$ est un Hamiltonien statique et $\mathcal{R}(t)$ une transformation unitaire dépendant du temps.

2. Soit $U(t)$ l'opérateur d'évolution associé à $H(t)$. On pose :

$$U(t) = \mathrm{e}^{\mathrm{i}\frac{\omega t}{2}\sigma_z}\,\tilde{U}(t) \ . \tag{18.291}$$

   Trouver l'équation différentielle satisfaite par $\tilde{U}$.

3. Intégrer cette équation (poser $\tilde{\omega} = \sqrt{[(\omega - \omega_0)/2]^2 + \Omega^2}$).

4. Mettre $U(t)$ sous la forme :

$$U(t) = \mathcal{R}(t)\,[u_0(t)\mathbf{1} + u_x(t)\sigma_x + u_z(t)\sigma_z] \tag{18.292}$$

   où les u_μ sont des scalaires.

5. En déduire l'expression exacte de $P_{12}(t)$. Analyser les différents cas :
   (a) $\omega \ll \omega_0$
   (b) $\omega = \omega_0$ (résonance, oscillation de Rabi)
   (c) $\omega \gg \omega_0$

6. Analyser le cas du "π-pulse", correspondant à une durée de perturbation T telle que $\omega T = \pi$.

18.7.12 Étude et mesure d'un spin $J = 1$

On considère un moment cinétique (spin) $\vec{J}$ ($J = 1$) dont les vecteurs propres communs à $\vec{J}^2$ et J_z sont notés simplement $|M\rangle$. Dans toute la suite, on raisonne exclusivement dans le sous-espace engendré par ces vecteurs. La base est ordonnée suivant les valeurs décroissantes de M.

1. Dans cette partie, on précise certaines relations utiles pour un tel moment cinétique.
   (a) Quelles sont les valeurs possibles de M ?
   (b) À quoi est égal $\vec{J}^2|M\rangle$?
   (c) Soit R la rotation dans l'espace $\mathbb{R}^3$ autour de la direction $\vec{n}$ d'un angle θ. À cette rotation est associé l'opérateur $\mathcal{R}$ agissant dans l'espace des états, dont l'expression est donnée en (17.98), le vecteur d'état transformé étant $|\Psi'\rangle = \mathcal{R}|\Psi\rangle$. A étant une observable, quelle est l'expression A' de l'observable transformée assurant l'invariance (galiléenne) des valeurs moyennes ?
   (d) Sans faire de calcul, préciser l'opérateur $\mathcal{R}_{zx}$ de la rotation qui transforme l'opérateur J_z en J_x.

(e) À l'aide des résultats précédents, compléter la relation suivante :

$$J_x = \mathrm{e}^{???}\, J_z\, \mathrm{e}^{???} \ . \tag{18.293}$$

(f) À quoi est égal $J_-|M\rangle$? En déduire la matrice de J_- et celle de J_+.

(g) Écrire les matrices 3×3 de J_x, J_y, J_z, J_x^2 et J_y^2.

(h) Exprimer J_z^3 en fonction de J_z. Qu'en est-il pour J_x^3 et J_y^3 ?

(i) Compte tenu du résultat obtenu en 1h, toute fonction de J_u ($u = x,\, y,\, z$) développable en série entière est égale à un polynôme du second degré en J_u. On pose :

$$\mathcal{R}_u(\theta) \stackrel{\text{déf}}{=} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\, J_u} = \mathbf{1} - \frac{\mathrm{i}}{\hbar} f(\theta) J_u + \frac{1}{\hbar^2}\, g(\theta) J_u^2 \ , \tag{18.294}$$

où f et g sont des fonctions à trouver.

i. Écrire une équation différentielle du *premier* ordre pour l'opérateur $\mathcal{R}_u(\theta)$

ii. En utilisant le résultat de 1h, écrire les équations différentielles pour les fonctions f et g.

iii. En déduire f, puis g.

(j) Soit $|M\rangle_x$ les états propres de J_x. L'équation aux valeurs propres pour J_x s'écrit :

$$J_x|M\rangle_x = M\hbar|M\rangle_x \ . \tag{18.295}$$

Quelles sont les valeurs possibles pour M ?

(k) En partant de (18.295) et en utilisant (18.293), montrer que $|M\rangle_x$ s'obtient en appliquant un certain opérateur de rotation (à préciser) au vecteur $|M\rangle$.

2. Dans cette partie, le moment cinétique est placé dans un champ magnétique auquel il se couple par son moment magnétique associé. Le Hamiltonien est $H = -\omega_{\mathrm{L}} J_z$.

(a) Justifier brièvement la forme de H.

(b) Quels sont les états propres et les valeurs propres de H ?

(c) L'état initial du spin étant $|M = 1\rangle$, quel est son état $|\Psi(t_1)\rangle$ à l'instant t_1 ?

(d) À l'instant t_1, on effectue une mesure de J_x et soit $P(M,\, t_1)$ la probabilité de trouver la valeur $M\hbar$ à cet instant. En utilisant le résultat 1k, compléter l'expression suivante de cette probabilité :

$$P(M,\, t_1) = \left|\Big(|M\rangle_?,\, \mathrm{e}^{-\frac{1}{\mathrm{i}\hbar}?J_?}|1\rangle\Big)\right|^? \ . \tag{18.296}$$

Dépend-elle de l'instant t_1 ?

(e) Utiliser la forme (18.294) pour expliciter (18.296), et mettre finalement cette probabilité sous la forme :

$$P(M,\, t_1) = \Big|\sum_{M'} A_{M'}\delta_{MM'}\Big|^2 \ . \tag{18.297}$$

Combien vaut $\sum_{M'} |A_{M'}|^2$?

(f) Le résultat de la première mesure ayant donné la valeur $M = 1$, quel est l'état du spin à $t = t_1 + 0$?

(g) On effectue une deuxième mesure de J_x à l'instant $t_2 = t_1 + T$ ($T > 0$). Expliquer pourquoi la probabilité de trouver la valeur M a pour expression :

$$P(M,\, t_2|+1,\, t_1) = \left|\left(|M\rangle_x,\, \mathrm{e}^{\mathrm{i}\omega_{\mathrm{L}}T\hbar^{-1}J_z}\mathrm{e}^{-\mathrm{i}\frac{\pi}{2}\hbar^{-1}J_y}|1\rangle\right)\right|^2 . \qquad (18.298)$$

(h) En déduire que $P(M,\, t_2|+1,\, t_1) = \left|\left(|M\rangle,\, \mathcal{R}_{zx}^{\dagger}\mathrm{e}^{\mathrm{i}\omega_{\mathrm{L}}T\hbar^{-1}J_z}\mathcal{R}_{zx}|1\rangle\right)\right|^2$.

(i) Montrer que :

$$\mathcal{R}_{zx}^{\dagger}\mathrm{e}^{\mathrm{i}\omega_{\mathrm{L}}T\hbar^{-1}J_z}\mathcal{R}_{zx} = \mathbf{1} - \frac{\mathrm{i}}{\hbar}\sin\omega_{\mathrm{c}}T\, J_x - \frac{1}{\hbar^2}(1-\cos\omega_{\mathrm{L}}T)\, J_x^2 . \qquad (18.299)$$

(j) En déduire $P(M,\, t_2|+1,\, t_1) \equiv p_M$ sous la forme :

$$P(M,\, t_2|+1,\, t_1) = \left|\sum_{M'}\tilde{A}_{M'}(t_2|t_1)\delta_{MM'}\right|^2 . \qquad (18.300)$$

Préciser les fonctions $\tilde{A}_{M'}(t_2|t_1)$.

(k) Donner les expressions de p_0, $p_{\pm 1}$.

(l) Interpréter ces résultats par une image gyroscopique.

18.7.13 À propos des polynômes de Legendre

1. En utilisant la fonction génératrice (18.110), montrer que :

$$P_l(X) = \sum_{n=0}^{[\frac{l}{2}]}(-1)^n\,\frac{(2l-2n)!}{2^l n!(l-n)!(l-2n)!}\,x^{l-2n} , \qquad (18.301)$$

où $[\frac{l}{2}]$ vaut $\frac{l}{2}$ si l est pair et $\frac{l-1}{2}$ si l est impair.

2. En utilisant $\frac{\mathrm{d}^n}{\mathrm{d}x^n}x^p = \frac{p!}{(p-n)!}\,x^{p-n}$, établir la formule de Rodrigues, (18.103).

3. À l'aide de la formule de Cauchy donnant l'expression intégrale de la l^{e} dérivée d'une fonction analytique, établir la représentation intégrale de Schäfli, (18.104).

4. Soit $f(x)$ une fonction intégrable sur $[-1,\, +1]$; utiliser la formule de Rodrigues pour montrer que :

$$\int_{-1}^{+1} P_l(x)f(x)\,\mathrm{d}x = \frac{1}{2^l l!}\int_{-1}^{+1}(1-x^2)^l\,\frac{\mathrm{d}^l}{\mathrm{d}x^l}f(x)\,\mathrm{d}x . \qquad (18.302)$$

5. En déduire la relation d'orthogonalité (18.112). On utilisera la fonction d'Euler de première espèce $B(p,\, q)$ définie comme $B(p,\, q) \stackrel{\text{déf}}{=} \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$, $\Gamma(z)$ étant la fonction d'Euler de deuxième espèce ; on a $\Gamma(n+1) = n!$ si $n \in \mathbb{N}$, et $\Gamma(\frac{1}{2}) = \sqrt{\pi}$; en outre ([38], § 8.335.1), $\Gamma(z)$ satisfait la formule de doublement $\sqrt{\pi}\,\Gamma(2z) = 2^{2z-1}\Gamma(z)(\Gamma(z+\frac{1}{2})$.

physique

Chapitre 19

Potentiel central et atome d'hydrogène

Ce chapitre est consacré au champ central ; après le traitement de quelques exemples, dont certains généralisent à trois dimensions des résultats obtenus dans $\mathbb{R}$, l'atome d'hydrogène est décrit en détail, avant une discussion approfondie de la symétrie dynamique spécifique du champ Coulombien.

19.1 Définition du champ central et exemples

Un champ de force conservatif dérive par définition d'un potentiel V, fonction qui dépend du rayon-vecteur fixant le point de l'espace où on calcule la force ; la relation est bien connue :

$$\vec{F}(\vec{r}) = -\vec{\nabla}V(\vec{r}) \ . \qquad (19.1)$$

Lorsque la force $\vec{F}$ est colinéaire à $\vec{r}$, elle est dite centrale ; alors, par le théorème du moment cinétique (voir (18.1)), le vecteur $\vec{L}$ est une constante du mouvement :

$$\boxed{\text{champ central} \iff \vec{F} \propto \vec{r} \iff \vec{L} \equiv \vec{r}\times\vec{p} = \mathrm{C^{ste}}} \qquad (19.2)$$

Dans ces conditions, seule la composante du gradient $\vec{\nabla}V$ parallèle à $\vec{r}$, égale à[1] $\partial_r V$, est non nulle. Il en résulte que l'énergie potentielle V ne dépend que de $r \stackrel{\text{déf}}{=} \|\vec{r}\|$, et se note

[1]On rappelle la convention d'écriture : $\partial_x \equiv \frac{\partial}{\partial x}$, etc.

plus simplement $V(r)$. Pour un champ central, et pour une particule de masse $m_{\rm p}$, le Hamiltonien s'écrit :

$$\boxed{H = \frac{\vec{p}^{\,2}}{2m_{\rm p}} + V(r)} \qquad (19.3)$$

Sauf mention explicite du contraire, on se place implicitement dans l'espace à trois dimensions $\mathbb{R}^3$.

En réalité, le problème se présente rarement d'emblée de cette façon. Souvent, on part de deux particules ponctuelles, sans structure interne, de masses m_1 et m_2 en interaction par une force obéissant à la loi de l'action et de la réaction : la force exercée par la particule 1 sur la particule 2 est opposée (vectoriellement) à la force exercée par la particule 2 sur la particule 1. En l'absence de toute direction privilégiée, notamment pour deux particules constituant un système isolé, l'énergie potentielle dépend alors de la seule distance entre les deux particules ; en outre, le mouvement du centre de masse est une translation uniforme. On peut alors se ramener à un problème plus simple : une fois évacuée la translation uniforme du centre de masse, subsiste un problème à trois degrés de liberté représentant le mouvement *relatif* des deux particules dans le repère de leur centre de masse, formellement identique à celui d'une particule unique de masse μ (masse réduite) égale à $m_1 m_2/(m_1 + m_2)$ et d'énergie potentielle $V(r)$; c'est ce que l'on appelle la "réduction du problème à deux corps en interaction centrale" (voir Tome I, sous-section 3.2.2).

L'exemple sans doute le plus connu est le potentiel Coulombien, d'ailleurs formellement identique au potentiel de gravitation[2], pour lequel $V(r)$ a la forme précise :

$$V(r) = \frac{q_1 q_2}{4\pi\varepsilon_0\, r} \ . \qquad (19.4)$$

Dans le cas d'un atome hydrogénoïde de charge nucléaire[3] $Z|e|$, l'énergie potentielle d'interaction entre le noyau et l'électron est :

$$V(r) = -\frac{Ze^2}{4\pi\varepsilon_0\, r} \equiv -\frac{Ze'^2}{r} \qquad \left(e'^2 = \frac{e^2}{4\pi\varepsilon_0}\right) \ . \qquad (19.5)$$

D'autres cas sont dignes d'intérêt : avec son énergie potentielle constante (en pratique nulle), la particule libre est un exemple de champ central, dont la trivialité n'est qu'apparente et fournit l'occasion d'une belle illustration des théorèmes de développement (on peut développer les ondes planes $\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$ sur les états propres à variables séparées – qui sont également propres du moment cinétique). Un autre exemple intéressant est celui du puits fini sphérique défini par :

$$V(r) = \begin{cases} -V_0 & \text{si } r < a \\ 0 & \text{si } r > a \end{cases} \ . \qquad (19.6)$$

[2]sauf que dans ce dernier cas les scalaires qui apparaissent (les masses) sont des quantités toujours essentiellement positives.

[3]e désigne partout la charge *négative* de l'électron : $e \simeq -1,6 \times 10^{-19}$ C.

Ces exemples seront abordés en temps utile, après formation explicite du Hamiltonien en coordonnées sphériques – coordonnées adaptées à la symétrie du problème considéré. Bien évidemment, on rencontrera ici et là des applications concrètes de résultats énoncés dans les deux chapitres précédents, à propos de la symétrie en général et du rôle du moment cinétique – il s'agira alors exclusivement du moment cinétique orbital.

19.1.1 Hamiltonien du problème central dans $\mathbb{R}^3$

Une fois réduit le problème à deux corps, le Hamiltonien du champ central dans $\mathbb{R}^3$ est :

$$H = \frac{\vec{p}^{\,2}}{2\mu} + V(r) \ , \qquad \vec{r} \in \mathbb{R}^3 \tag{19.7}$$

μ étant la masse réduite. Conformément à la prescription standard de la représentation-q, $\vec{p} = -\mathrm{i}\hbar\vec{\nabla} = -\mathrm{i}\hbar(\partial_x, \partial_y, \partial_z)$; dès lors, pour obtenir l'expression de l'énergie cinétique en coordonnées sphériques, deux méthodes sont possibles :

- la première consiste à effectuer le changement de variables dans les opérateurs différentiels représentant l'énergie cinétique T, pour passer des coordonnées cartésiennes (x, y, z) aux coordonnées sphériques (r, θ, ϕ). Un calcul facile mais laborieux donne le résultat suivant :

$$\vec{p}^{\,2} = -\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\Big(r^2\,\frac{\partial}{\partial r}\Big) - \frac{\hbar^2}{r^2}\Big(\frac{\partial^2}{\partial \theta^2} + \cot\theta\,\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\Big) \ . \tag{19.8}$$

  On reconnaît dans la partie angulaire l'expression du carré du moment cinétique obtenue au chapitre 18, éq. (18.58) ; d'où l'expression de l'énergie cinétique T :

$$T \stackrel{\text{déf}}{=} \frac{\vec{p}^{\,2}}{2\mu} = -\frac{\hbar^2}{2\mu}\frac{1}{r^2}\,\frac{\partial}{\partial r}\Big(r^2\,\frac{\partial}{\partial r}\Big) + \frac{\vec{L}^{\,2}}{2\mu r^2} \ . \tag{19.9}$$

  Le terme $\vec{L}^{\,2}/(2\mu r^2)$ est appelé terme centrifuge : il apparaît dans l'équation fondamentale de la dynamique pour la variable radiale r et, reporté au second membre, ressort comme une force d'inertie complémentaire ; comme il s'agit de rotation, c'est bien un terme centrifuge.

- la deuxième méthode part des expressions classiques ; on a :

$$\vec{L}^{\,2} = (\vec{r}\times\vec{p})^2 = r^2p^2\,\sin^2\Theta = r^2p^2\,(1-\cos^2\Theta) \equiv \vec{r}^{\,2}\vec{p}^{\,2} \ - \ (\vec{r}.\vec{p})^2 \ , \tag{19.10}$$

  Θ étant l'angle entre $\vec{r}$ et $\vec{p}$, d'où l'on tire :

$$\vec{p}^{\,2} = \Big(\frac{\vec{r}.\vec{p}}{r}\Big)^2 + \frac{\vec{L}^{\,2}}{r^2} \ . \tag{19.11}$$

  Cette expression classique doit être convenablement symétrisée afin d'engendrer *in fine* un opérateur hermitique après la substitution $\vec{p} \ \rightarrow \ -\mathrm{i}\hbar\vec{\nabla}$. Il n'y a pas

d'ambiguïté pour le rapport $\vec{L}^2/r^2$ car les deux opérateurs commutent entre eux ($\vec{L}$ est associé aux rotations, qui ne changent pas r). La seule difficulté réside dans le terme p_r^2 :

$$p_r^2 \overset{\text{déf}}{=} \left(\frac{\vec{r}.\vec{p}}{r}\right)^2 \quad \Longleftrightarrow \quad p_r = \frac{\vec{r}.\vec{p}}{r} \ , \tag{19.12}$$

dont la symétrisation naturelle est :

$$p_r = \frac{1}{2}\left(\frac{\vec{r}}{r}.\vec{p} + \vec{p}.\frac{\vec{r}}{r}\right) \ . \tag{19.13}$$

En vertu de (19.14), on peut écrire[4] :

$$\vec{p}.\frac{\vec{r}}{r} = [\vec{p}, \frac{\vec{r}}{r}] + \frac{\vec{r}}{r}.\vec{p} = -\mathrm{i}\hbar\left(\vec{\nabla}.\frac{\vec{r}}{r}\right) + \frac{\vec{r}}{r}.\vec{p} \ ; \tag{19.15}$$

l'expression symétrisée (19.13) conduit alors à :

$$p_r = \frac{\vec{r}}{r}.\vec{p} - \frac{\mathrm{i}\hbar}{2}\left(\vec{\nabla}.\frac{\vec{r}}{r}\right) = -\mathrm{i}\hbar\left[\frac{\vec{r}}{r}.\vec{\nabla} + \frac{1}{2}\left(\vec{\nabla}.\frac{\vec{r}}{r}\right)\right] \ . \tag{19.16}$$

Le premier terme entre crochets contient la composante du gradient le long du rayon-vecteur : c'est simplement ∂_r. Le deuxième terme se calcule facilement à partir de sa définition :

$$\left(\vec{\nabla}.\frac{\vec{r}}{r}\right) = \frac{\partial}{\partial x}\frac{x}{r} + \frac{\partial}{\partial y}\frac{y}{r} + \frac{\partial}{\partial z}\frac{z}{r} \ . \tag{19.17}$$

Le terme en x est $\partial_x \frac{x}{r} = \frac{1}{r} - \frac{x^2}{r^3}$; en additionnant les trois termes en x, y et z, il vient donc :

$$\left(\vec{\nabla}.\frac{\vec{r}}{r}\right) = \frac{3}{r} - \frac{x^2+y^2+z^2}{r^3} = \frac{2}{r} \ , \tag{19.18}$$

d'où, selon (19.16) :

$$\boxed{p_r = -\mathrm{i}\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right) \equiv -\mathrm{i}\hbar\,\frac{1}{r}\frac{\partial}{\partial r}r} \tag{19.19}$$

On remarque au passage que p_r n'est pas égal à $-\mathrm{i}\hbar\partial_r$ en conséquence du fait que r n'est pas une coordonnée rectangulaire (cartésienne). Avec l'expression (19.19), on trouve toujours $[r, p_r] = \mathrm{i}\hbar$, mais cette relation n'a rien d'évident *a priori* puisque précisément les coordonnées ne sont pas cartésiennes[5].

[4]Pour toute fonction vectorielle $\vec{f}(\vec{r})$ on a :

$$[\vec{p}, \vec{f}(\vec{r})] = -\mathrm{i}\hbar\,(\vec{\nabla}.\vec{f}(\vec{r}))\mathbf{1} \ ; \tag{19.14}$$

les parenthèses sont là pour signifier que l'opérateur gradient, au second membre, n'agit que sur $\vec{f}(\vec{r})$, pas sur une fonction sous-entendue à droite.

[5]D'ailleurs, toute expression du genre $\varpi(r) \equiv -\mathrm{i}\hbar[\partial_r + \phi(r)]$ conduit elle aussi à la relation de commutation $[r, \varpi(r)] = \mathrm{i}\hbar$.

Il reste à calculer le carré de p_r ; en prenant garde à l'ordre des opérateurs dans l'élévation au carré du binôme :

$$\Big(\frac{\partial}{\partial r}+\frac{1}{r}\Big)^2=\frac{\partial}{\partial r}\frac{\partial}{\partial r}+\frac{\partial}{\partial r}\frac{1}{r}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\ , \tag{19.20}$$

on trouve[6] :

$$\boxed{p_r^2=-\hbar^2\Big(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\Big)\equiv-\hbar^2\,\frac{1}{r^2}\,\frac{\partial}{\partial r}\Big(r^2\,\frac{\partial}{\partial r}\Big)} \tag{19.22}$$

En ajoutant $\vec{L}^2/(2\mu r^2)$, l'expression (19.9) est retrouvée.

En définitive, le Hamiltonien du problème central dans $\mathbb{R}^3$ s'écrit :

$$\boxed{H=\frac{p_r^2}{2\mu}+\frac{\vec{L}^2}{2\mu r^2}+V(r)\equiv-\,\frac{\hbar^2}{2\mu}\,\frac{1}{r^2}\frac{\partial}{\partial r}\,r^2\,\frac{\partial}{\partial r}+\frac{\vec{L}^2}{2\mu r^2}+V(r)\equiv T+V} \tag{19.23}$$

où chaque opérateur différentiel agit sur *tout* ce qui se trouve à sa droite. T est l'opérateur représentant l'énergie cinétique de la particule.

Pour la référence ultérieure, il est utile de préciser les conditions à satisfaire pour que cet opérateur soit hermitique. Le carré du moment cinétique commute avec r, donc le second terme de (19.23) est visiblement hermitique ; $V(r)$ l'est aussi. La seule question porte sur le premier terme, qui sera hermitique si p_r l'est. Comme il s'agit d'un opérateur différentiel, il faut que toutes les fonctions sur lequel il agit se comportent convenablement aux bornes $r=0$ et $r=+\infty$ (la condition d'hermiticité se vérifie par des intégrations par parties et il suffit bien que les termes tout intégrés soient nuls).

Soit $\psi_1(r)$ et $\psi_2(r)$ deux fonctions quelconques ; la condition d'hermiticité de p_r s'écrit par définition :

$$\langle\psi_1|p_r\psi_2\rangle=\langle p_r\psi_1|\psi_2\rangle\ , \tag{19.24}$$

soit :

$$\int_0^{+\infty}\psi_1^*(r)\big[p_r\psi_2(r)\big]\,r^2\mathrm{d}r=\int_0^{+\infty}\big[p_r\psi_1(r)\big]^*\psi_2(r)\,r^2\mathrm{d}r\ . \tag{19.25}$$

Compte tenu de (19.19), cette égalité s'explicite en :

$$-\mathrm{i}\hbar\int_0^{+\infty}\psi_1^*(r)\Big[\Big(\frac{\partial}{\partial r}+\frac{1}{r}\Big)\psi_2(r)\Big]r^2\mathrm{d}r=+\mathrm{i}\hbar\int_0^{+\infty}\Big[\Big(\frac{\partial}{\partial r}+\frac{1}{r}\Big)\psi_1^*(r)\Big]\psi_2(r)\,r^2\mathrm{d}r\ . \tag{19.26}$$

---

[6] D'une façon générale, dans $\mathbb{R}^D$, p_r^2 est donné par (voir exercice 19.3.1) :

$$p_r^2=-\hbar^2\Big(\frac{\partial^2}{\partial r^2}+\frac{D-1}{r}\frac{\partial}{\partial r}+\frac{(D-1)(D-3)}{4r^2}\Big)\ . \tag{19.21}$$

Noter la spécificité exclusive des dimensions $D=1$ et $D=3$, pour lesquelles le terme en $\frac{1}{r^2}$ est absent. Dans tous les autres cas, ce terme joue le rôle d'une barrière répulsive floue lorsque $r\to 0$, contraignant la fonction radiale à être toujours nulle à l'origine, quel que soit le moment cinétique.

Le premier membre est :

$$-\mathrm{i}\hbar \int_0^{+\infty} \psi_1^*(r)\left[\psi_2'(r) + r^{-1}\psi_2(r)\right] r^2 \mathrm{d}r \quad . \tag{19.27}$$

Le second membre se transforme par intégration par parties ; on trouve :

$$+\mathrm{i}\hbar \left[r^2\psi_1^*(r)\psi_2(r)\right]_0^{+\infty} - \mathrm{i}\hbar \int_0^{+\infty} \psi_1^*(r)\left[r\,\psi_2(r) + r^2\psi_2'(r)\right] \mathrm{d}r \quad . \tag{19.28}$$

En comparant (19.27) et (19.28), on voit que la condition d'hermiticité (19.26) s'écrit :

$$\left[r^2\psi_1^*(r)\psi_2(r)\right]_0^{+\infty} = 0 \ ; \tag{19.29}$$

en particulier, pour tout élément diagonal, il faut :

$$\left[r^2|\psi(r)|^2\right]_0^{+\infty} = 0 \ . \tag{19.30}$$

Pour que ces conditions soient satisfaites, il est *suffisant* que toutes les fonctions de l'espace vectoriel satisfassent les égalités suivantes :

$$\boxed{\lim_{r\rightarrow 0}\left[r\,|\psi(r)|\right] = 0 \qquad \lim_{r\rightarrow +\infty}\left[r\,|\psi(r)|\right] = 0} \tag{19.31}$$

Ces conditions suffisantes, notamment celle à l'infini, sont en général satisfaites pour les états liés[7], bien que la condition de normalisabilité, à elle seule, ne suffise pas à assurer l'hermiticité de p_r. En effet, si par exemple ψ se comporte comme $r^{-\alpha}$ à l'origine, $r^2|\psi|^2$ est sommable en zéro si $\alpha < \frac{3}{2}$; avec cette seule dernière restriction, $r\psi(r)$ peut violer la condition (19.31).

♦ *Remarque*

L'hermiticité de p_r assure celle de H. Toutefois, p_r n'est pas une observable au sens strict[8] : ses fonctions propres n'appartiennent pas à l'espace de fonctions qui sont à la fois de carré sommable et telles que $[r\psi(r)]_{r=0} = 0$. En effet, d'après (19.19) l'équation propre pour p_r est :

$$-\mathrm{i}\hbar\frac{1}{r}\frac{\partial}{\partial r}[rf(r)] = \lambda\, f(r) \ . \tag{19.32}$$

En multipliant membre à membre par r, on voit que les solutions sont les ondes sphériques :

$$f(r) = \frac{\mathrm{C^{ste}}}{r}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\lambda r} \ . \tag{19.33}$$

Ces fonctions ne sont pas de carré sommable et ne satisfont pas la condition suffisante assurant que p_r est hermitique. On ne peut donc faire jouer les postulats de

[7]Pour un état lié, donc normalisable, il est *nécessaire* (dans $\mathbb{R}^3$) que $\lim_{r\rightarrow+\infty}\left[r^2\,|\psi(r)|^2\right] = 0$.
[8]voir [34], p. 293.

la Mécanique quantique à propos de p_r puisque ses états propres sont en-dehors de l'espace sur lequel cet opérateur est hermitique[9]. De toute façon, c'est l'hermiticité de p_r^2 qui est nécessaire (et suffisante) ; il n'est d'ailleurs pas évident que l'hermiticité de p_r^2 exige celle de p_r – étant entendu que si un opérateur est hermitique, son carré l'est aussi. ♦

Le Hamiltonien H (19.23) possède visiblement la symétrie sphérique ; on est donc assuré d'avance que H et n'importe laquelle des composantes de $\vec{L}$ commutent :

$$\boxed{[H, \vec{L}\,] = 0} \tag{19.34}$$

Techniquement, ceci résulte d'une part du fait que $\vec{L}$ et T commutent, puisque l'ordre des deux opérations : rotation, dérivation par rapport à r, est sans importance (géométriquement, l'indifférence à l'ordre est évidente). D'autre part, $\vec{L}$ commute avec $V(r)$: quand on fait tourner le système, par définition d'un champ central, l'énergie potentielle ne change pas. L'équation (19.34) assure que l'on peut trouver des états propres communs à H, $\vec{L}^2$ et L_z. Si l'on note $\psi(r,\theta,\phi)$ ces derniers, on voit que, compte tenu de la forme[10] de H, tous les états propres de ce dernier opérateur peuvent s'obtenir comme des combinaisons linéaires[11] de fonctions à variables séparées :

$$\boxed{\psi(r,\theta,\phi) = R(r)\,Y_{lm}(\theta,\phi)} \tag{19.35}$$

où les Y_{lm} sont les harmoniques sphériques introduites dans la section 18.3 ; reportant une telle forme dans l'équation aux valeurs propres $H\psi \equiv (T+V)\psi = E\psi$, il vient :

$$\Big[-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \frac{\vec{L}^2}{2\mu\, r^2} + V(r)\Big]\,R(r)\,Y_{lm}(\theta,\phi) = ER(r)\,Y_{lm}(\theta,\phi) \ . \tag{19.36}$$

La fonction $R(r)$ est une constante pour l'opérateur différentiel $\vec{L}^2$; en faisant agir celui-ci sur ses états propres Y_{lm}, et après simplification[12], il reste l'équation dite radiale pour la seule fonction $R(r)$:

$$-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\Big(r^2\frac{\mathrm{d}}{\mathrm{d}r}\Big)R(r) + \frac{\hbar^2 l(l+1)}{2\mu r^2}\,R(r) + V(r)\,R(r) = ER(r) \ . \tag{19.37}$$

[9]Cette propriété de p_r n'est pas si particulière que cela : p_x présente la même anomalie puisque les ondes planes ne sont pas normalisables. Il n'empêche que la condition d'hermiticité est tout autant requise pour p_x que pour p_r.

[10]H est une combinaison linéaire de deux opérateurs – l'un radial, l'autre angulaire – où les coefficients multiplicatifs dépendent au plus des variables autres que celles concernées par l'opérateur.

[11]En raison de la symétrie de rotation autour notamment de Oz, tous les états propres du type (19.35), de même l et différant par leur m, ont la même énergie. Toute combinaison linéaire de ces états est donc encore propre, avec la même énergie.

[12]L'action explicite de $\vec{L}^2$ fait apparaître le simple produit $\hbar^2 l(l+1)Y_{lm}$, de sorte que l'harmonique sphérique se trouve en facteur dans l'équation complète et peut être omise dans la suite.

Il est souvent avantageux d'introduire une autre fonction inconnue[13] $u(r) \stackrel{\text{déf}}{=} rR(r)$; par substitution dans (19.37), on voit que $u(r)$ satisfait l'équation suivante :

$$\boxed{-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2u}{\mathrm{d}r^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r)\right]u(r) = E\,u(r)} \tag{19.38}$$

soit :

$$H_{\text{eff}}(r) \stackrel{\text{déf}}{=} -\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2u}{\mathrm{d}r^2} + V_{\text{eff}}(r)\,u(r) = E\,u(r) \ , \quad V_{\text{eff}}(r) \stackrel{\text{déf}}{=} V(r) + \frac{\hbar^2 l(l+1)}{2\mu r^2} \ . \tag{19.39}$$

On voit ainsi apparaître un Hamiltonien effectif H_{eff}, paramétré par l'entier l, dont l'équation aux valeurs et fonctions propres est nettement plus simple que (19.37), et présente de plus l'avantage d'être formellement un problème à une dimension (mais réduit à $\mathbb{R}_+$) avec un potentiel effectif incorporant le terme centrifuge (ce dernier est nul pour un moment cinétique nul, ce qui n'est pas surprenant). Ainsi, le problème pour $u(r)$ est exactement le même que celui d'une particule à une dimension, d'énergie potentielle $V_{\text{eff}}(x)$ et en outre confinée sur le demi-axe réel positif par une barrière infranchissable en $x = 0$. Tout ce que l'on sait à propos du mouvement à une dimension (réalité des fonctions d'onde, dégénérescence, etc., voir Tome I, chapitre 15) peut ainsi être utilisé, tant que la fonction $u(r)$ est seule pertinente. En particulier[14], si $u_{l,\,\alpha}(r)$ est une fonction propre de (19.38), α désignant le(s) nombre(s) quantique(s) complétant la spécification cet état propre, deux fonctions propres $u_{l,\,\alpha}(r)$ et $u_{l,\,\alpha'}(r)$ sont orthogonales si $\alpha \neq \alpha'$; toutefois, pour deux valeurs distinctes $l \neq l'$, aucune propriété d'orthogonalité n'est à attendre puisque deux solutions $u_{l,\,\alpha}(r)$ et $u_{l',\,\alpha'}(r)$ sont propres de deux Hamiltoniens effectifs *différents*.

Si la limite $\lim_{r\to+\infty} V(r) \stackrel{\text{déf}}{=} V_\infty$ est infinie (par exemple : oscillateur harmonique à trois dimensions), tous les états sont liés. Si V_∞ est fini, le seuil en énergie séparant états liés et non-liés est $E = V_\infty$ – sauf cas exceptionnel : on connaît en effet quelques cas exotiques où il existe des états liés dont l'énergie est noyée dans le *continuum* d'états de diffusion [35], une prévision de von Neumann et Wigner [36] dès 1929. En tout état de cause, la condition d'hermiticité de p_r – qui assure celle de H – exige :

$$\lim_{r\to 0}\left[rR(r)\right] = 0 \ , \qquad \lim_{r\to\infty}\left[rR(r)\right] = 0 \ , \tag{19.40}$$

soit :

$$\boxed{\lim_{r\to 0} u(r) = 0 \qquad \lim_{r\to\infty} u(r) = 0} \tag{19.41}$$

La sous-section 19.1.2 donne une analyse détaillée du comportement de la fonction radiale en $r = 0$, $r = +\infty$ et au voisinage d'un saut de potentiel.

[13] Le préfacteur r entre les fonctions $R(r)$ et $u(r)$ est lié à la dimension de l'espace (ici $D = 3$). En effet, à trois dimensions, la condition de normalisation s'écrit $\int_{\mathbb{R}^3} r^2|R(r)|^2\,\mathrm{d}r = 1$, soit $\int_0^{+\infty}|u(r)|^2\,\mathrm{d}r = 1$. Dans $\mathbb{R}^D$, la normalisation est $\int_{\mathbb{R}^D} r^{D-1}|R(r)|^2\,\mathrm{d}r = 1$, de sorte que posant $u(r) = r^{(D-1)/2}R(r)$ on retrouve la même condition pour $u(r)$, soit $\int_0^{+\infty}|u(r)|^2\,\mathrm{d}r = 1$.

[14] Le plus souvent, la dépendance paramétrique des solutions vis-à-vis de l sera sous-entendue ; elle n'est ici explicite que pour préciser le point en discussion.

La fonction radiale R est essentiellement réelle, comme le montre d'abord l'analyse de l'équation de conservation :

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{j} = 0 \ , \qquad \rho \overset{\text{déf}}{=} \Psi^* \Psi \ , \qquad \vec{j} \overset{\text{déf}}{=} \frac{\hbar}{2\mathrm{i}\mu} (\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*) \tag{19.42}$$

En effet, quand Ψ est un état stationnaire (donc de la forme $\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Et} RY_{lm}$), la densité ρ est constante en temps et il reste $\operatorname{div} \vec{j} = 0$. Comme les angles sont découplés de r et sans interaction entre eux, les trois composantes du courant, j_r, j_θ et j_ϕ, sont séparément des constantes dans l'espace ; la valeur de ces constantes dépend de la nature de l'état, lié ou non.

Pour un état lié, seule la composante j_ϕ peut être différente de zéro (elle ne l'est pas forcément). La raison à ceci est d'ordre purement topologique : l'angle ϕ est la seule coordonnée qui se "boucle" dans le système, permettant des courants permanents non nuls. Au contraire, l'angle θ (qui varie entre 0 et π) n'autorise pas de circulation permanente, et il en va de même pour la coordonnée radiale r, qui va de 0 à $+\infty$. En particulier, la composante radiale j_r a pour expression :

$$j_r(r, \theta, \phi) = \frac{\hbar}{2\mathrm{i}\mu} |Y_{lm}(\theta, \phi)|^2 \left(R^* \frac{\mathrm{d}R}{\mathrm{d}r} - R \frac{\mathrm{d}R^*}{\mathrm{d}r} \right) \ ; \tag{19.43}$$

comme elle est nulle pour un état lié, on a :

$$R^* \frac{\mathrm{d}R}{\mathrm{d}r} = R \frac{\mathrm{d}R^*}{\mathrm{d}r} \quad \Longleftrightarrow \quad \ln R = \ln R^* + \mathrm{C}^{\text{ste}} \quad \Longleftrightarrow \quad R^* \propto R \ . \tag{19.44}$$

R et R^* sont deux fonctions proportionnelles et représentent donc le même état physique.

La réalité de R peut se voir d'une autre façon, en invoquant l'absence de dégénérescence et la symétrie par renversement du temps[15]. La fonction $u(r)$ introduite ci-dessus satisfait de fait l'équation à une dimension sur $\mathbb{R}_+$ (19.38) ; partant de là, on peut établir un théorème du Wronskien, comme on le fait pour les problèmes à une dimension sur $\mathbb{R}$ (voir Tome I, section 15.1). On trouve ainsi :

$$\big[W(u_1, u_2)\big]_a^b \equiv \big[u_1 u_2' - u_2 u_1'\big]_a^b = (E_1 - E_2) \int_a^b u_1(r) u_2(r) \, \mathrm{d}r \ . \tag{19.45}$$

En particulier, si u_1 et u_2 sont associées à la *même* valeur propre E ($E_1 = E_2 = E$), la variation du Wronskien entre a et b est nulle :

$$\big[W(u_1, u_2)\big]_a^b = 0 \ , \tag{19.46}$$

et puisque les deux valeurs a et b sont quelconques, on en déduit :

$$W(u_1, u_2) = \mathrm{C}^{\text{ste}} \ . \tag{19.47}$$

[15] Cette même symétrie permet aussi de se convaincre que Y_{lm} et Y_{l-m} ont la même énergie – indépendamment de la symétrie de rotation autour de Oz.

Ceci permet d'établir que les états discrets (i.e. normalisables) $u(r)$ sont non-dégénérés. En effet, pour de tels états, les fonctions u_i et u_i' sont nécessairement nulles à l'infini, le Wronskien est donc nul à l'infini. Comme il prend la même valeur partout, il est nul partout. Il vient ainsi, $\forall\, r$:

$$u_1(r)u_2'(r) - u_2(r)u_1'(r) = 0 \quad \Longleftrightarrow \quad u_2(r) \propto u_1(r) \ . \tag{19.48}$$

Ainsi, deux fonctions u_1 et u_2 associées à la même valeur propre E sont proportionnelles : à chaque E correspond de fait une seule fonction propre[16] u. Compte tenu de la symétrie par renversement du temps, u et u^* correspondent à une et une seule valeur propre ; deux telles fonctions diffèrent donc au plus par une phase globale, dénuée de sens physique : en définitive, la fonction radiale R peut bien toujours être prise réelle. Dans ces conditions, les Y_{lm} étant toujours conventionnellement normalisées par rapport aux variables angulaires, la normalisation de la fonction d'onde d'un état lié impose la normalisation à part de $R(r)$:

$$\int_0^{+\infty} R^2(r)\, r^2 \mathrm{d}r = 1 \quad \Longleftrightarrow \quad \int_0^{+\infty} u^2(r)\, \mathrm{d}r = 1 \ . \tag{19.49}$$

En ce qui concerne le mouvement radial non-lié, il n'est pas non plus dégénéré (pour un l donné). Dans $\mathbb{R}$, tout mouvement non-lié est dégénéré deux fois, puisque l'inversion de la vitesse ne change pas l'énergie (encore une manifestation de l'invariance par renversement du temps). En revanche, l'introduction d'une barrière parfaitement réfléchissante brise la symétrie gauche-droite, et sur $\mathbb{R}_+$ le mouvement non-lié perd ainsi sa dégénérescence ; il en va de même ici. La fonction $u(r)$ doit s'annuler en $r = 0$, ce qui montre que même dans le cas libre, on n'a pas la liberté de choisir une combinaison linéaire *arbitraire* de deux solutions linéairement indépendantes (comme ce serait le cas en présence d'une dégénérescence égale à 2) : force est de prendre la seule et unique fonction nulle à l'origine, qui est bien essentiellement réelle (pour plus de détails, voir sous-section 19.1.3). L'identité avec le problème unidimensionnel permet de bien comprendre le rôle de la barrière en $r = 0$: comme tout ce qui arrive repart, le courant est forcément nul.

19.1.2 Comportements de la fonction radiale

D'après la discussion précédente, la fonction d'onde du champ central peut être recherchée sous la forme ($R(r) = u(r)/r$) :

$$\boxed{\psi(r, \theta, \phi) = \frac{u(r)}{r}\, Y_{lm}(\theta, \phi)} \tag{19.50}$$

[16]Le spectre de H reste bien sûr en général dégénéré : n et l étant fixés, tous les états propres distincts de L_z ont la même énergie en l'absence de direction privilégiée. En outre, le phénomène de dégénérescence "accidentelle" (E ne dépendant en fait que de n) reste toujours possible. Pour le champ Coulombien, cette dégénérescence particulière donne la même énergie à toutes les fonctions radiales de même n : R_{nl} et $R_{nl'}$ ont la même énergie (voir section 19.2.3). Il n'y a pas de théorème du Wronskien pour un tel couple : la démonstration ci-dessus suppose que les deux fonctions u_1 et u_2 sont propres du *même* Hamiltonien ; or le Hamiltonien central (19.38) dépend de l par le terme centrifuge.

où la seule fonction encore inconnue $u(r)$ satisfait l'équation (19.38). Comme r ne change pas dans l'inversion d'espace, la fonction ψ a la même parité que Y_{lm}, soit $(-1)^l$. Les états propres sont donc pairs (*resp.* impairs) si l est un entier pair (*resp.* impair).

Tout état lié doit être normalisable ; les Y_{lm} étant supposées normalisées à part, il convient seulement d'assurer que $u(r)$ est une fonction de module carré sommable selon (19.49), ce qui élimine une grande partie des solutions mathématiques de l'équation différentielle (19.39). En outre, cette équation est vraie partout sauf évidemment en $r = 0$; la singularité du point $r = 0$ doit donc faire l'objet d'une prescription spéciale, qui a déjà été établie en arguant du fait que p_r est hermitique (voir (19.41)), assurant que le Hamiltonien l'est aussi.

La discussion générale des conditions que doit satisfaire toute solution $u(r)$ physiquement acceptable est sans intérêt ; pour les potentiels possédant à la fois des états liés et des états non-liés[17], l'immense majorité des cas pratiques est couverte quand on fait les hypothèses additionnelles suivantes sur la fonction $V(r)$:

$$V(r) \rightarrow 0 \quad \text{si} \quad r \rightarrow +\infty \ , \tag{19.51}$$

$$V(r) \simeq A\,r^{\alpha} \quad \text{si} \quad r \rightarrow 0 \ , \tag{19.52}$$

où $\alpha \geq -1$. La discussion couvre donc des potentiels divergents à l'origine et en particulier le cas du champ coulombien[18].

Examinons d'abord précisément ce qui se passe à l'origine, lorsque l'hypothèse (19.52) est satisfaite. Partant de l'équation pour u, l'intégration membre à membre dans un voisinage à droite donne :

$$\int_0^{\delta r} \Big[-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2 u}{\mathrm{d}r^2} + \Big(\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r)\Big)\,u(r)\Big]\,\mathrm{d}r = E\int_0^{\delta r} u(r)\,\mathrm{d}r \ . \tag{19.53}$$

Le second membre est nul à la limite $\delta r \rightarrow 0$ puisque $\lim_{r\rightarrow 0} u(r) = 0$; après transformation évidente du premier membre, on doit avoir :

$$\lim_{\delta r \rightarrow 0}\Big\{-\frac{\hbar^2}{2\mu}\Big[\frac{\mathrm{d}u}{\mathrm{d}r}\Big]_0^{\delta r} + \int_0^{\delta r}\Big[\frac{\hbar^2 l(l+1)}{2\mu r^2} + A\,r^{\alpha}\Big]\,u(r)\Big\}\,\mathrm{d}r = 0 \ . \tag{19.54}$$

Faisons maintenant l'hypothèse que u se comporte comme r^{β} dans le voisinage de l'origine (comme u doit tendre vers zéro, ceci impose $\beta > 0$) ; dans cette hypothèse, la quantité entre accolades dans (19.54) se transforme en :

$$-\frac{\hbar^2}{2\mu}\Big[\beta r^{\beta-1}\Big]_0^{\delta r} + \int_0^{\delta r}\Big[\frac{\hbar^2 l(l+1)}{2\mu}\,r^{\beta-2} + A r^{\alpha+\beta}\Big]\,\mathrm{d}r \ . \tag{19.55}$$

[17] L'existence d'états non-liés est assurée par le fait que $V(r)$ a une limite *finie* quand $r \rightarrow +\infty$. Ce n'est pas le cas pour l'oscillateur harmonique à trois dimensions qui, comme son homologue sur $\mathbb{R}$, ne possède que des états liés.

[18] Toutefois, la discussion suivante ne considère donc pas le cas où $V(r)$ contient un terme en r^{-2}, du même type que le potentiel centrifuge. Ce cas est examiné dans le problème 19.3.15.

À condition que β soit différent de 1, il vient, après intégration et regroupement des termes :

$$-\frac{\hbar^2}{2\mu}\left[\left(\beta-\frac{l(l+1)}{\beta-1}\right)r^{\beta-1}\right]_0^{\delta r}+\left[\frac{A}{\alpha+\beta+1}\,r^{\alpha+\beta+1}\right]_0^{\delta r} . \qquad (19.56)$$

Comme $\alpha+1\geq 0$ et $\beta>0$, le second terme tend vers zéro dans tous les cas : il faut donc $\beta>1$, auquel cas R est en $r^{\beta-1}$ et tend vers zéro. Dans le cas où $\beta=1$, $u\sim Cr$, l'intégration dans (19.55) fournit un terme logarithmique et la condition (19.54) s'écrit :

$$\lim_{\delta r\to 0}\left\{\left[\frac{\hbar^2}{2\mu}\,l(l+1)\,\ln r+\frac{A}{\alpha+2}\,r^{\alpha+2}\right]_0^{\delta r}\right\}=0 . \qquad (19.57)$$

Sauf si $l=0$, cette condition ne peut pas être satisfaite à cause du logarithme ; en conséquence, seuls les états sphériques (notés "s") peuvent se comporter à l'origine comme Cr (pour u) et donc comme une constante (pour R).

L'analyse précédente repose sur une intégration locale et est donc sûre ; elle montre bien que u doit tendre vers zéro à l'origine – mais ne permet pas de trouver l'exposant β. Pour l'obtenir, il suffit de manipuler l'équation différentielle elle-même, avec toujours l'hypothèse $u\propto r^{\beta}$. Il vient :

$$-\frac{\hbar^2}{2\mu}\,\beta(\beta-1)\,r^{\beta-2}+\frac{\hbar^2 l(l+1)}{2\mu}\,r^{\beta-2}+A\,r^{\alpha+\beta}=Er^{\beta} ; \qquad (19.58)$$

β étant positif et $\alpha\geq -1$, on a $\alpha+\beta\geq -1$; l'annulation des termes les plus divergents donne :

$$\beta(\beta-1)=l(l+1) \quad\Longleftrightarrow\quad \beta=l+1 . \qquad (19.59)$$

En définitive, en présence d'un potentiel $V(r)\propto r^{\alpha}$, $(\alpha\geq -1)$, u se comporte comme r^{l+1} à l'origine : l étant positif ou nul, la fonction u s'annule bien toujours en $r=0$, conformément à la comparaison faite avec un mouvement sur $\mathbb{R}_+$ confiné par une barrière parfaitement réfléchissante. En conséquence, la fonction radiale $R(r)$ varie comme r^l ; seuls les états s ont une densité de probabilité de présence non nulle en $r=0$ ($R(r)\propto r^0$) :

$$\boxed{r\sim 0\ :\ V(r)\sim r^{\alpha},\ \alpha\geq -1 \quad\Longrightarrow\quad R(r)\sim r^l \Longleftrightarrow u(r)\sim r^{l+1}} \qquad (19.60)$$

Discutons maintenant brièvement le comportement de $R(r)$ à l'infini, en se cantonnant aux états liés ($E<0$ compte tenu de (19.51)). On voit tout de suite que u ne peut avoir un comportement en pure loi-puissance ($u\propto r^{-\lambda}$, $\lambda>0$) à l'infini, car il serait impossible de satisfaire l'équation (19.38) avec les termes dominants. En revanche, ceci devient possible si u contient un facteur exponentiel ; on a ainsi :

$$\boxed{r\to+\infty\ :\ R(r)\sim r^{\lambda}\,\mathrm{e}^{-kr} \qquad (k=\sqrt{-2\mu E/\hbar^2}\,)} \qquad (19.61)$$

où la valeur précise de l'exposant reste à trouver – comme il dépend du comportement asymptotique précis de $V(r)$ à l'infini, il ne peut être obtenu par une discussion générale ; le point important à retenir est que, pour un état lié, $R(r)$ a un comportement essentiellement exponentiel à l'infini.

Examinons enfin ce qui se passe lorsque $V(r)$ présente un saut fini pour une certaine valeur r_0. L'intégration de l'équation pour u de part et d'autre de cette singularité fournit :

$$-\frac{\hbar^2}{2\mu}\Big[u'(r_0+\delta r)-u'(r_0-\delta r)\Big]+\int_{r_0-\delta r}^{r_0+\delta r}\Big[\frac{\hbar^2 l(l+1)}{2\mu r^2}+V(r)\Big]u(r)\,\mathrm{d}r = E\int_{r_0-\delta r}^{r_0+\delta r}u(r)\,\mathrm{d}r\ . \tag{19.62}$$

Comme $|u|^2$ est une densité de probabilité, $|u|^2$ doit être partout localement sommable. En admettant (au pire) pour u une divergence comme $(r-r_0)^{-\lambda}$, il faut que $\lambda < 1/2$ et la limite $\delta r = 0$ de l'intégrale du second membre est nulle. Dans l'intégrale au premier membre, le terme en $l(l+1)/r^2$ est constant près de r_0 et $u(r)$ est sommable. Quant au terme en $V(r)u(r)$, il est donc en $(r-r_0)^{-\lambda}\,\theta(r-r_0)$, avec $\lambda < 1/2$: son intégrale est nulle. Finalement, on obtient :

$$\boxed{\lim_{\delta r\to 0}[u'(r_0+\delta r)-u'(r_0-\delta r)]=0} \tag{19.63}$$

une égalité qui exprime la continuité de la dérivée et, par voie de conséquence, celle de $u(r)$. La fonction radiale est continue et à dérivée continue même en présence d'un saut *fini* de potentiel ; ceci n'est pas surprenant puisque, formellement, l'équation aux valeurs propres pour u est identique à celle d'un mouvement sur $\mathbb{R}_+$.

19.1.3 La particule libre

Une particule libre possède une énergie potentielle constante, que l'on peut toujours prendre égale à zéro. Ce problème – dépourvu d'états liés pour des raisons physiques évidentes – peut évidemment se résoudre en coordonnées cartésiennes ; alors les états propres de H sont des ondes planes :

$$\boxed{\langle\vec{r}\,|\vec{k}\rangle \stackrel{\text{déf}}{=} \psi_{\vec{k}}(\vec{r}) = C\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}} \tag{19.64}$$

où $\vec{k}$ est un vecteur de composantes réelles, faute de quoi $\psi_{\vec{k}}$ divergerait exponentiellement à l'infini. Ces fonctions ne sont pas normalisables, mais servent de base, au moyen de l'intégrale de Fourier, à la construction de paquets d'ondes normalisables, non-stationnaires mais solutions de l'équation de Schrödinger, un état initial étant prescrit. L'énergie associée à la fonction propre (19.64) est reliée au vecteur d'onde $\vec{k}$ par la relation habituelle :

$$\boxed{E(\vec{k}) = \frac{\hbar^2}{2\mu}\,\vec{k}^{\,2}} \tag{19.65}$$

qui constitue la relation de dispersion. Chaque état est infiniment dégénéré, puisque l'énergie ne dépend que du module k de $\vec{k}$, pas de son orientation (isotropie de l'espace). Il n'y a pas de quantification de l'énergie, toutes les valeurs de $k \geq 0$ étant autorisées.

Les états propres sont orthogonaux au sens où[19] :

$$\int_{\mathbb{R}^3} \mathrm{d}^3r\,\psi^*_{\vec{k}}(\vec{r})\,\psi_{\vec{k}\,'}(\vec{r}) = 0 \quad \forall \vec{k}\,' \neq \vec{k}\ . \tag{19.69}$$

Si la constante C dans (19.64) est choisie égale[20] à $(2\pi)^{-3/2}$, la relation d'orthonormalisation des états propres s'écrit :

$$\langle \vec{k}\,|\vec{k}\,'\rangle \equiv \int_{\mathbb{R}^3} \psi^*_{\vec{k}}(\vec{r})\,\psi_{\vec{k}\,'}(\vec{r})\,\mathrm{d}^3r = \delta(\vec{k} - \vec{k}\,')\ , \tag{19.70}$$

tandis que la relation de fermeture est :

$$\int_{\mathbb{R}^3} \langle \vec{r}\,|\vec{k}\rangle\langle \vec{k}\,|\vec{r}\,'\rangle\,\mathrm{d}^3k \equiv \int_{\mathbb{R}^3} \psi_{\vec{k}}(\vec{r})\,\psi^*_{\vec{k}}(\vec{r}\,')\,\mathrm{d}^3k = \delta(\vec{r} - \vec{r}\,')\ . \tag{19.71}$$

Les problèmes de non-normalisabilité des ondes planes peuvent aussi se régler sur la base d'arguments purement physiques. Il est toujours loisible de choisir dans un premier temps une grande boîte cubique de quantification, de coté L, permettant de manipuler uniquement des états discrets, moyennant le choix judicieux de conditions aux limites ; la longueur L peut être aussi grande que l'on veut, mais elle est *finie* (la taille de la pièce de manip', la taille de notre galaxie,...). Dès lors, aucun infini ne peut surgir en cours de route. En principe, la limite $L \longrightarrow +\infty$ prise en fin de calcul restitue les mêmes résultats que le calcul direct fait en manipulant des objets quelque peu symboliques[21]. La représentation en ondes planes correspond à un ensemble complet d'observables qui commutent (ECOC) constitué des trois composantes de l'impulsion et de l'énergie.

Bien évidemment, rien n'interdit de résoudre le problème de la particule libre en coordonnées sphériques : le potentiel nul est visiblement un cas particulier de potentiel à symétrie sphérique ; ce choix étant fait, on obtient une autre représentation de l'état

---

[19] Schématiquement, ces résultats proviennent de :

$$I_L(k) \stackrel{\text{déf}}{=} \int_{-L/2}^{+L/2} \mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}x = \frac{2}{k}\sin\frac{kL}{2} \tag{19.66}$$

si $k \neq 0$, avec trivialement $I_L(0) = L$. Connaissant l'intégrale de Dirichlet $\int_{-\infty}^{+\infty} \frac{\sin x}{x}\mathrm{d}x = \pi$, on obtient dans la limite $L \longrightarrow +\infty$ la représentation standard de la fonction de Dirac :

$$\delta(k) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}kx}\mathrm{d}x\ . \tag{19.67}$$

Lorsque l'argument de δ est un vecteur, il faut comprendre que, x, y et z étant les coordonnées *cartésiennes* de $\vec{r}$, on a :

$$\boxed{\delta(\vec{r}) \stackrel{\text{déf}}{=} \delta(x)\delta(y)\delta(z)} \tag{19.68}$$

[20] auquel cas $\psi_{\vec{k}}$ est un nombre (sans dimension physique).

[21] Dans le cas contraire, le problème traité aurait, pour des raisons à élucider, une sensibilité à des conditions aux limites non pertinentes physiquement.

Il convient cependant de garder à l'esprit que l'introduction des bonnes échelles pour éliminer les divergences et régulariser la théorie ne doit pas dispenser de garantir d'un bout à l'autre le respect d'éventuelles contraintes topologiques.

quantique, correspondant *ipso facto* à un ECOC constitué cette fois de ($\vec{L}^2$, L_z et H). Comme $\vec{p}$ et $\vec{L}$ ne commutent pas, ces deux représentations sont incompatibles, au sens des observables incompatibles. Comme précédemment, et pour les mêmes raisons, on pose la factorisation (19.35) ; la seule fonction inconnue, $R(r)$, satisfait :

$$-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\Big(r^2\frac{\mathrm{d}}{\mathrm{d}r}\Big)R(r)+\frac{\hbar^2 l(l+1)}{2\mu r^2}R(r)=ER(r)\ . \tag{19.72}$$

En posant :

$$k=\sqrt{2\mu E/\hbar^2}\in\mathbb{R}_+\ ,\qquad \rho=kr\ ,\qquad R(r)=v(\rho)\ , \tag{19.73}$$

on trouve l'équation pour la fonction v :

$$\frac{\mathrm{d}^2v}{\mathrm{d}\rho^2}+\frac{2}{\rho}\frac{\mathrm{d}v}{\mathrm{d}\rho}+\Big[1-\frac{l(l+1)}{\rho^2}\Big]v(\rho)=0\ . \tag{19.74}$$

À un détail près (le facteur 2 dans le terme en $\rho^{-1}v'$), cette équation est très voisine de celle définissant les fonctions de Bessel J_ν, laquelle s'écrit :

$$J_\nu''(z)+\frac{1}{z}J_\nu'(z)+\Big(1-\frac{\nu^2}{z^2}\Big)J_\nu(z)=0\ . \tag{19.75}$$

J_ν admet le développement en série entière :

$$J_\nu(z)=\Big(\frac{z}{2}\Big)^\nu\sum_{n=0}^{+\infty}\frac{(-1)^n}{n!\Gamma(\nu+n+1)}\Big(\frac{z}{2}\Big)^{2n}\qquad(-\pi<\arg z<\pi)\ , \tag{19.76}$$

où $\Gamma(\xi)$ est la fonction d'Euler de deuxième espèce (qui coïncide avec la factorielle pour $\xi\in\mathbb{N}^*$: $\Gamma(n+1)=n!$). Ceci s'établit facilement en posant que $J_\nu(z)$ est de la forme $z^\beta\sum_{n\in\mathbb{N}}c_nz^n$ et en reportant dans l'équation différentielle (19.75) ; on trouve d'abord $\beta=\nu$, puis la relation de récurrence entre les coefficients c_n, le tout reconstituant le développement (19.76). Pour $\nu\notin\mathbb{Z}$, $J_\nu(z)$ a un point de branchement à l'origine – d'où la précision nécessaire sur $\arg z$ dans (19.76) ; dans le cas contraire où ν est entier, J_ν est une fonction entière pour laquelle cette condition est superfétatoire.

Afin d'obtenir strictement (19.75) à partir de (19.74), posons $v(\rho)\stackrel{\text{déf}}{=}\rho^\alpha f(\rho)$; le report dans (19.74) donne :

$$\frac{\mathrm{d}^2f}{\mathrm{d}\rho^2}+\frac{2(\alpha+1)}{\rho}\frac{\mathrm{d}f}{\mathrm{d}\rho}+\Big[1-\frac{l(l+1)-\alpha(\alpha+1)}{\rho^2}\Big]f(\rho)=0\ ; \tag{19.77}$$

il suffit maintenant de choisir $\alpha=-\frac{1}{2}$ pour obtenir :

$$\frac{\mathrm{d}^2f}{\mathrm{d}\rho^2}+\frac{1}{\rho}\frac{\mathrm{d}f}{\mathrm{d}\rho}+\Big[1-\frac{(l+\frac{1}{2})^2}{\rho^2}\Big]f(\rho)=0\ , \tag{19.78}$$

qui montre par comparaison avec (19.75) que les deux fonctions $f(\rho)$ et $J_{\pm(l+\frac{1}{2})}$, satisfaisant la même équation différentielle homogène, sont proportionnelles l'une à l'autre. Sur le développement (19.76), on voit que $J_\nu(\rho)$ est le produit de $(\rho/2)^\nu$ et d'une série

entière représentant une fonction analytique dans $\mathbb{C}$ (fonction *entière*). Ici, on a *a priori* le choix $\nu = \pm(l+\frac{1}{2})$, mais comme il faut physiquement que $R(r) \sim r^l$ quand $r \sim 0$, soit $\rho^{-1/2}f(\rho) \sim \rho^l$, la seule possibilité est $\nu = +(l+\frac{1}{2})$; en définitive, la fonction radiale est $R(r) \propto r^{-1/2}J_{l+1/2}$. Il est d'usage d'introduire les fonctions de Bessel *sphériques* $j_l(r)$, définies à partir des fonctions des J_ν ([37], p. 437) :

$$\boxed{j_l(\rho) \stackrel{\text{déf}}{=} \sqrt{\frac{\pi}{2\rho}}\, J_{l+\frac{1}{2}}(\rho)} \tag{19.79}$$

Le préfacteur contenant $\rho^{-1/2}$ fait de chaque j_l une fonction entière (alors que J_ν est en général branchée à l'origine). En utilisant l'une des représentations intégrales connues pour J_ν, on voit que les j_l peuvent notamment s'exprimer comme suit :

$$j_l(\rho) = \frac{\rho^l}{2^{l+1}\, l!} \int_{-1}^{+1} \mathrm{e}^{\mathrm{i}\rho s}\, (1-s^2)^l\, \mathrm{d}s \ . \tag{19.80}$$

L'intérêt de cette représentation est de mettre en évidence le lien entre les j_l et les polynômes de Legendre déjà rencontrés (voir chapitre 18, section 18.3). En effet, en effectuant l intégrations par parties, on trouve :

$$j_l(\rho) = \mathrm{i}^{-l}\, \frac{(-1)^l}{2^{l+1}\, l!} \int_{-1}^{+1} \mathrm{e}^{\mathrm{i}\rho s}\, \frac{\mathrm{d}^l}{\mathrm{d}s^l}(1-s^2)^l\, \mathrm{d}s \ . \tag{19.81}$$

Si l'on se souvient alors de la définition des polynômes de Legendre obtenus lors de l'étude des harmoniques sphériques, on réalise le lien étroit entre le polynôme de degré l, P_l, et la fonction de Bessel sphérique j_l ; de fait, compte tenu de la formule de Rodrigues (18.103) et de sa conséquence (18.104), on a :

$$\boxed{j_l(\rho) = \frac{1}{2\,\mathrm{i}^l} \int_{-1}^{+1} \mathrm{e}^{\mathrm{i}\rho s} P_l(s)\, \mathrm{d}s} \tag{19.82}$$

Il est clair sur cette dernière expression que $j_l(\rho)$ est une fonction à valeurs réelles, conformément aux arguments généraux développés p. 785. Les expressions des premières fonctions j_l sont :

$$j_0(\rho) = \frac{\sin\rho}{\rho} \ , \quad j_1(\rho) = \frac{\sin\rho}{\rho^2} - \frac{\cos\rho}{\rho} \ , \quad j_2(\rho) = \left(\frac{3}{\rho^3} - \frac{1}{\rho}\right) \sin\rho - \frac{3}{\rho^2}\cos\rho \ . \tag{19.83}$$

La solution de l'équation radiale pour la particule libre est donc de la forme :

$$\boxed{R_{El}(r) = C_{El}\, j_l(kr)} \tag{19.84}$$

où la constante C_{El} est déterminée par une convention de normalisation, par exemple[22] :

$$C_{El}\, C_{E'l} \int_0^{+\infty} j_l(kr)\, j_l(k'r)\, r^2 \mathrm{d}r = \delta(E-E') \ , \tag{19.85}$$

[22] Il n'y a pas lieu d'attendre l'orthogonalité de j_l et de $j_{l'}$ $(l \neq l')$; en effet, ces deux fonctions sont propres de deux Hamiltoniens *différents*, comme le montre l'équation (19.74), qui est paramétrée par l.

auquel cas la dimension de $R_{El}(r)$ est (longueur$^3\times$énergie)$^{-1/2}$. Une normalisation alternative fréquente est $R_{kl}(r) = C_{kl}\, j_l(kr)$, la constante C_{kl} étant cette fois fixée par :

$$C_{kl}\, C_{k'l} \int_0^{+\infty} j_l(kr)\, j_l(k'r)\, r^2 \mathrm{d}r = \delta(k-k') \ , \tag{19.86}$$

et alors $[R_{kl}(r)] = \text{longueur}^{-1}$.

Au contraire des ondes planes, le module des fonctions $j_l(kr)$ n'est pas constant et décroît même vers zéro en oscillant ; toutefois, les fonctions ne sont pas non plus normalisables car :

$$\int_0^{+\infty} |j_l(kr)|^2\, r^2 \mathrm{d}r = +\infty \ ; \tag{19.87}$$

en effet, une méthode de col [13] permet de montrer que $J_\nu(\rho) \sim \sqrt{\pi/(2\rho)}\, \cos(\rho - \nu\frac{\pi}{2} - \frac{\pi}{4})$ quand $|\rho| \gg |\nu|$; il en résulte que, à un facteur oscillant près, $j_l(kr) \sim (kr)^{-1}$ quand $kr \gg 1$, et n'est donc pas de module carré sommable. Enfin, avec (19.80) (ou (19.76)), on voit que $j_l(\rho) \propto \rho^l$ quand $\rho \ll 1$ (alors, au plus bas ordre en ρ, l'intégrale est une constante), assurant que $R_{El}(r)$ est bien en r^l quand $r \to 0$.

À ce stade, on dispose finalement de deux jeux de fonctions propres pour la particule libre d'énergie donnée E : les ondes planes (19.64) ayant cette énergie, en nombre infini non-dénombrable (toutes les orientations possibles pour $\vec{k}$), et les fonctions $j_l(kr)Y_{lm}(\theta, \phi)$, de même énergie, qui forment au contraire un ensemble infini dénombrable ($l \in \mathbb{N}$, $-l \leq m \leq +l$). Il s'agit toutefois dans les deux cas d'un système complet de fonctions et on peut donc développer n'importe quelle fonction d'un jeu en série des fonctions de l'autre jeu. Évidemment, la première fonction étant choisie et paramétrée par k – qui représente l'énergie E –, elle se développe sur tous les états de l'autre jeu ayant la *même* énergie E :

$$\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} = \sum_{l=0}^{+\infty} \sum_{m=-l}^{+l} c_{lm} j_l(kr) Y_{lm}(\theta, \phi) \ , \tag{19.88}$$

où l'on retrouve le *même* module k du vecteur $\vec{k}$ dans les deux membres. Prenons en particulier le cas où le vecteur $\vec{k}$ est dirigé le long de Oz ; alors le premier membre ne dépend que de θ, plus de ϕ, et seules les harmoniques sphériques Y_{l0} apparaissent dans le développement ; on se souvient que l'harmonique sphérique Y_{l0} est proportionnelle au polynôme de Legendre $P_l(\cos\theta)$ (voir (18.122)), de sorte que l'on peut écrire :

$$\mathrm{e}^{\mathrm{i}kr\cos\theta} = \sum_{l=0}^{+\infty} c_l j_l(kr) P_l(\cos\theta) \ . \tag{19.89}$$

Compte tenu de la relation d'orthogonalité des polynômes P_l (voir (18.112)), les coefficients c_l ont une expression simple. En effet, multipliant membre à membre (19.89) par $P_l(X = \cos\theta)$ et en intégrant sur X, on obtient :

$$c_l j_l(kr) = \frac{2l+1}{2} \int_{-1}^{+1} \mathrm{e}^{\mathrm{i}krX}\, P_l(X)\, \mathrm{d}X \ ; \tag{19.90}$$

par comparaison avec l'expression intégrale (19.82), on voit que $c_l = (2l+1)\mathrm{i}^l$, d'où finalement *l'importante égalité suivante, particulièrement utile dans les problèmes de* diffusion (voir chapitre 26, tout particulièrement section 26.4) :

$$\boxed{\mathrm{e}^{\,ikr\cos\theta} = \sum_{l=0}^{+\infty}(2l+1)\,\mathrm{i}^l\,j_l(kr)P_l(\cos\theta)} \tag{19.91}$$

Évidemment, il est possible d'effectuer les développements inverses et de trouver les coefficients $d_{lm}(k)$ tels que :

$$j_l(kr)Y_{lm}(\theta,\phi) = \int_0^{\pi}\sin\Theta\mathrm{d}\Theta\int_0^{2\pi}\mathrm{d}\Phi\,d_{lm}(k)\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\ . \tag{19.92}$$

où Θ et Φ sont les angles fixant la direction des vecteurs $\vec{k}$ ayant tous le même module k (paramètre qui est, à nouveau, présent aux deux membres et caractérise l'énergie commune des états tous dégénérés que l'on développe les uns sur les autres). Tout comme la base d'ondes planes, la base propre sphérique j_lY_{lm} permet également de représenter n'importe quel paquet d'ondes, en particulier un paquet gaussien (voir problème 19.3.6).

19.1.4 Puits "carré" sphérique

Il s'agit ici de résoudre[23] l'équation radiale avec l'énergie potentielle $V(r)$:

$$V(r) = \begin{cases} -V_0 & \text{si } r < a \\ 0 & \text{si } r > a \end{cases} \qquad (V_0 > 0)\ . \tag{19.93}$$

$V(r)$ a une variation "carrée", mais, physiquement, le puits a évidemment une symétrie sphérique. L'équation aux valeurs propres radiale est :

$$\Big[-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\Big(r^2\frac{\mathrm{d}}{\mathrm{d}r}\Big)+\frac{\hbar^2 l(l+1)}{2\mu r^2}\Big]R(r) = \begin{cases} (E+V_0)R(r) & \text{si } r<a \\ ER(r) & \text{si } r>a \end{cases}\ , \tag{19.94}$$

Considérons d'abord les états liés, pour lesquels $-V_0 < E < 0$, (voir Tome I, éq. (15.24)). Dans la région $r < a$, on a :

$$-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\Big[\Big(r^2\frac{\mathrm{d}}{\mathrm{d}r}\Big)R(r)\Big]+\frac{\hbar^2 l(l+1)}{2\mu r^2}R(r) = (E+V_0)R(r)\ . \tag{19.95}$$

Au changement $E \rightarrow E+V_0 > 0$, c'est l'équation (19.72) ; posant cette fois $E+V_0 = \frac{\hbar^2K^2}{2\mu}$, $K > 0$, on peut donc immédiatement écrire (voir (19.84)) :

$$r < a\ :\qquad R_{El}(r) = C_{El}\,j_l(Kr)\ . \tag{19.96}$$

[23] [34], p. 304.

Dans l'autre région, $r > a$, on retrouve à nouveau la même équation (19.72), mais le facteur au second membre, l'énergie E, est maintenant négatif, ce qui signifie que, pour r grand, la fonction $R(r)$ satisfait :

$$-\frac{\hbar^2}{2\mu}\,R''(r)\;\simeq\;E\,R(r) \equiv -|E|\,R(r)\ . \tag{19.97}$$

R est donc essentiellement[24] une combinaison d'exponentielles réelles ; comme toujours, seules les solutions non-divergentes sont acceptables. Ce point étant précisé, la résolution de l'équation se fait presque sans calculs ; en effet, rien n'interdit de continuer à poser $E = \frac{\hbar^2 k^2}{2\mu}$, mais comme $E < 0$, k est maintenant imaginaire pur :

$$k = \mathrm{i}\kappa\ , \qquad \kappa = \sqrt{\frac{2\mu(-E)}{\hbar^2}}\ . \tag{19.98}$$

En posant encore $\rho = kr$ et $R(r) = v(\rho)$ en guise d'intermédiaire, l'équation pour $R(r)$ est encore (19.74) ; avec $v(\rho) = \rho^{-1/2}f(\rho)$, on retrouve (19.78). Finalement, posant $\rho = \mathrm{i}\sigma$, $f(\rho) = g(\sigma)$, soit $\sigma = \kappa r \in \mathbb{R}$, l'équation pour g s'écrit :

$$\frac{\mathrm{d}^2 g}{\mathrm{d}\sigma^2} + \frac{1}{\sigma}\frac{\mathrm{d}g}{\mathrm{d}\sigma} - \left[1 + \frac{(l+\frac{1}{2})^2}{\sigma^2}\right] g(\sigma) = 0\ . \tag{19.99}$$

Cette équation possède deux solutions linéairement indépendantes[25] traditionnellement notées [38] I_ν et K_ν (avec $\nu = l + \frac{1}{2}$) ; la première diverge exponentiellement à l'infini, et doit donc être rejetée. La fonction K_ν (fonction de Hankel), au contraire, décroît exponentiellement à l'infini et constitue donc la seule solution acceptable physiquement. En définitive[26] :

$$r > a\ : \qquad R_{El}(r) = C'_{El}\, r^{-1/2}\, K_{l+\frac{1}{2}}(\kappa r)\ , \qquad \kappa^2 = -\frac{2\mu E}{\hbar^2} > 0\ . \tag{19.100}$$

Pour les états sphériques ($l = 0$), on obtient la fonction $K_{1/2}(\sigma) = \sqrt{\frac{\pi}{2\sigma}}\mathrm{e}^{-\sigma}$, soit :

$$r > a\ : \qquad R_{El=0}(r) = C'_{E0}\,\frac{1}{r}\,\mathrm{e}^{-\kappa r}\ , \tag{19.101}$$

qui sont des ondes sphériques amorties. Au total, les fonction propres sont oscillantes (les j_l) à l'intérieur du puits, et décroissent (en gros) exponentiellement à l'extérieur de celui-ci, ces deux aspects étant typiques des états liés.

À l'instar du problème unidimensionnel, et comme montré plus haut (voir 19.63)), il convient d'écrire les équations de continuité (raccordement) de R et de sa dérivée[27]

[24] i.e. à un préfacteur algébrique r^λ près.

[25] Ces fonctions appartiennent visiblement à la classe des fonctions de Bessel, mais leur argument est imaginaire pur d'où, pour la clarté (et par usage), des notations spécifiques rappelant cette propriété. Par exemple, on a : $I_\nu(z) \stackrel{\text{déf}}{=} \mathrm{e}^{-\mathrm{i}\nu\frac{\pi}{2}}\, J_\nu(\mathrm{i}z)$.

[26] On définit aussi des fonctions (sphériques) h_l reliées à $K_{l+1/2}$ tout comme les j_l sont reliées aux $J_{l+1/2}$.

La notation K_ν est traditionnelle : ne pas faire de confusion avec $K \stackrel{\text{déf}}{=} \sqrt{2\mu(E+V_0)/\hbar^2}$.

[27] Il est équivalent d'écrire la continuité de la dérivée logarithmique R'/R.

en $r = a$ (la condition de normalisabilité, évacuant d'emblée les solutions divergentes à l'infini, est déjà prise en compte), les relations qui en découlent conduisent finalement à une équation dont les zéros fournissent les seules valeurs possibles de k, donc de l'énergie. Une fois encore, les conditions physiques requises pour la fonction d'onde produisent la quantification spontanée de l'énergie des états liés.

Les conditions de raccordement produisent une équation dépendant de la valeur du moment cinétique orbital[28], donnée par l'entier l. Pour les états s ($l = 0$), ces conditions s'expriment très simplement, puisqu'alors[29] (voir éq. (19.83)) :

$$R_{E\,l=0}(r) = \begin{cases} C_{E0}\frac{\sin Kr}{r} & \text{si } r < a \\ C'_{E0}\frac{1}{r}\,\mathrm{e}^{-kr} & \text{si } r > a \end{cases} . \tag{19.102}$$

Les conditions de continuité de $u(r)$ et de $u'(r)$ pour $l = 0$ sont :

$$u(a-0) = u(a+0) \quad \Longleftrightarrow \quad C_{E0}\sin Ka = C'_{E0}\mathrm{e}^{-ka} \tag{19.103}$$

et :

$$u'(a-0) = u'(a+0) \quad \Longleftrightarrow \quad C_{E0}K\cos Ka = -kC'_{E0}\mathrm{e}^{-ka} . \tag{19.104}$$

En égalant les deux expressions du rapport $\frac{C'_{E0}}{C_{E0}}$, on obtient :

$$ka = -\frac{Ka}{\tan Ka} \quad \Longleftrightarrow \quad \sqrt{k_0^2 - K^2}\,a = -\frac{Ka}{\tan Ka} \qquad (k_0^2 = 2\mu V_0/\hbar^2) , \tag{19.105}$$

avec $K \leq k_0$ puisque l'énergie E est négative mais bornée inférieurement par $-V_0$ (voir Tome I, éq. (15.24)). Cette équation est réminiscente de ce qui a été obtenu à une dimension d'espace (voir problème 15.6.13) ; ce n'est pas surprenant : pour $l = 0$, l'équation pour $u(r)$ – fonction qui doit être continue et à dérivée continue – est identique à l'équation aux valeurs propres pour le même potentiel à $D = 1$ (complété avec la barrière en $r = 0$), puisque le terme centrifuge est nul dans le cas considéré ($l = 0$). Pour les états s, on obtient donc toujours, à trois dimensions, les mêmes résultats (pour l'énergie et la fonction $u(r)$) que pour une particule soumise au même potentiel mais confinée sur le demi-axe réel positif par une barrière infranchissable située à l'origine : $u(r)$ est donc $\propto \sin Kr$, et $R_0(r) \propto \frac{\sin Kr}{r}$. Il n'est guère surprenant de retrouver pour les états *sphériques* la même dépendance en r que pour les états unidimensionnels par rapport à x : en un sens, un état à symétrie sphérique est quelque peu unidimensionnel.

Le puits carré, qui a toujours au moins un état lié sur tout $\mathbb{R}$ (voir Tome I, sous-section 15.3.1), perd cette propriété si une telle barrière est présente : celle-ci a un effet répulsif[30], donc déstabilisant et il n'y a d'état lié que si le puits est assez profond et pas

[28] voir ci-dessous le cas des états non-liés, éqs. (19.115) et (19.113), et aussi [39], problème 63, p. 153.

[29] On note que $R_{E0}(r = 0) \neq 0$, conformément à ce qui a été dit plus haut pour les états s.

[30] De la même façon, un puits $V(x) = g\delta(x - x_0)$, $g < 0$, $x_0 > 0$ offre toujours un et un seul état lié à toute particule de masse m_p ; si on ajoute une barrière réfléchissante (par exemple en $x = 0$), l'état lié ne subsiste que si $|g| > \hbar^2/(2m_\mathrm{p}x_0)$: si la barrière est trop proche du puits (voir Tome I, problème 15.6.8 et [40]) et/ou la particule est trop légère, il n'y a plus de piégeage possible (voir aussi le problème 15.6.13).

trop étroit. Il en va de même ici : il est facile de voir que l'équation ci-dessus n'a de solution que si $k_0 a > \pi/2$, c'est-à-dire :

$$k_0 a > \frac{\pi}{2} \quad \Longleftrightarrow \quad V_0 > \frac{\pi^2\hbar^2}{8\mu a^2} \ . \tag{19.106}$$

Sans surprise, le nombre d'états liés N_b augmente quand V_0 augmente (avec a constant), très précisément d'une unité à chaque fois que k_0 franchit une valeur du genre $\frac{\pi}{2} + n\pi$, n entier naturel. Inversement, à V_0 constant, N_b décroît si la portée a diminue ; au-dessous de la valeur-seuil déduite de (19.106) :

$$\boxed{a_< \stackrel{\text{déf}}{=} \left(\frac{\pi^2\hbar^2}{8\mu V_0}\right)^{1/2}} \tag{19.107}$$

il n'y a plus d'état lié. Un puits sphérique de profondeur donnée mais trop étroit ne peut piéger une particule, et ce d'autant moins que la masse de celle-ci est petite.

À l'inverse, le puits infiniment profond contient une infinité d'états liés. Considérant d'abord les états sphériques $l = 0$ dont le vecteur d'onde K est donné par (19.105), il suffit d'examiner la limite de cette équation quand $k_0 \to +\infty$, la quantité $\frac{\hbar^2 K^2}{2\mu} \equiv E + V_0$ donnant l'énergie des états liés comptée à partir du fond du puits. Géométriquement, cette équation est l'intersection du quart de cercle $\sqrt{d^2 - x^2}$ ($d = k_0 a$, $x = Ka$) avec le graphe de la courbe $-\frac{x}{\tan x}$. On voit immédiatement que si le rayon du quart de cercle augmente indéfiniment, les abscisses des points d'intersection se rapprochent des multiples positifs de π. En définitive, les énergies des états sphériques du puits infiniment profond sont :

$$\boxed{E_{n\,l=0}^{(\infty)} = n^2\frac{\pi^2\hbar^2}{2\mu a^2} \qquad (n \in \mathbb{N}^*)} \tag{19.108}$$

quand l'origine de l'énergie est prise au fond du puits. Ce résultat était prévisible : comme déjà dit, l'équation radiale est, pour les états s, strictement la même que pour le puits carré unidimensionnel – et il s'agit dans cette limite du puits infiniment profond localisé entre les abscisses 0 et a. Les fonctions propres sont donc $u_{n\,0}(r) \propto \sin n\pi\frac{r}{a}$ pour $0 \leq r \leq a$ (et nulles ailleurs), soit :

$$\boxed{R_{n\,l=0}^{(\infty)}(r) = \sqrt{\frac{2}{a}}\,\frac{1}{r}\sin\frac{n\pi r}{a} \qquad (0 \leq r \leq a\,,\, n \in \mathbb{N}^*)} \tag{19.109}$$

qui sont donc des fonctions sphériques $j_0(n\pi r/a)$, convenablement normalisées.

Pour les états $l > 0$, et toujours en comptant l'énergie $E = \frac{\hbar^2 K^2}{2\mu}$ à partir du fond du puits, les fonctions propres sont à nouveau des $j_l(Kr)$ (voir (19.96)), mais elles sont maintenant contraintes par $j_l(Ka) = 0$, puisque la particule ne peut sortir du puits et que les fonctions propres doivent être continues. Cette condition de raccordement quantifie les valeurs de K, dont les valeurs possibles, à l fixé, sont ainsi données par $K_{nl} = \beta_{nl}a^{-1}$, où β_{nl} est le n^e zéro de la fonction de Bessel sphérique $j_l(\beta)$.

Pour terminer, et revenant au puits de profondeur finie, déterminons les états non-liés, l'énergie E étant maintenant positive ; avec la paramétrisation naturelle :

$$E = \frac{\hbar^2 k^2}{2\mu} \ , k > 0 \ , \qquad E + V_0 = \frac{\hbar^2 K^2}{2\mu} \ , \qquad K^2 = k^2 + k_0^2 > 0 \ . \tag{19.110}$$

Lorsque $r < a$, la solution est encore donnée par l'expression (19.96) :

$$r < a \ : \qquad R_{El}(r) = C_{El}\, r^{-1/2}\, J_{l+\frac{1}{2}}(Kr) \ . \tag{19.111}$$

Pour $r > a$, $R_l(r)$ satisfait l'équation obtenue pour la particule libre (19.72) ; avec la paramétrisation (19.110), et en introduisant la fonction $v(\rho) \stackrel{\text{déf}}{=} \rho^{-1/2} f(\rho)$, on retrouve l'équation différentielle (19.78), dont les solutions sont des fonctions de Bessel $J_{\pm(l+1/2)}$. Ici apparaît une différence majeure : il n'y a plus maintenant aucune raison de disqualifier la solution $J_{-(l+1/2)}(kr)$, puisque $r > a$ (pour $r < a$, elle avait été rejetée car divergente en $r = 0$) ; exprimée à l'aide des fonctions de Bessel, la solution la plus générale à l'extérieur du puits est donc la combinaison linéaire quelconque :

$$r > a \ : \qquad R_{El}(r) = r^{-1/2}[C_{E+}\, J_{l+\frac{1}{2}}(kr) + C_{E-}\, J_{-(l+\frac{1}{2})}(kr)] \ , \tag{19.112}$$

où les deux constantes $C_{E\pm}$ sont pour l'instant arbitraires et où, comme on va le voir, k n'est astreint à aucune condition et peut prendre toutes les valeurs dans $\mathbb{R}_+$. Pour déterminer $C_{E\pm}$, il faut et suffit d'écrire les conditions de raccordement exprimant la continuité de $R_l(r)$ et de sa dérivée en $r = a$. On doit ainsi avoir :

$$C_{El}\, J_{l+\frac{1}{2}}(Ka) = C_{E+}\, J_{l+\frac{1}{2}}(ka) + C_{E-}\, J_{-(l+\frac{1}{2})}(ka) \ , \tag{19.113}$$

et :

$$C_{El}\Big[-\frac{1}{2a} J_{l+\frac{1}{2}}(Ka) + K J'_{l+\frac{1}{2}}(Ka)\Big] = -\frac{1}{2a}[C_{E+} J_{l+\frac{1}{2}}(ka) + C_{E-} J_{-(l+\frac{1}{2})}(ka)] + k[C_{E+}\, J'_{l+\frac{1}{2}}(ka) + C_{E-}\, J'_{-(l+\frac{1}{2})}(ka)] \ ; \tag{19.114}$$

compte tenu de (19.113), cette dernière condition se simplifie en :

$$C_{El}\, \frac{K}{k}\, J'_{l+\frac{1}{2}}(Ka) = C_{E+}\, J'_{l+\frac{1}{2}}(ka) + C_{E-}\, J'_{-(l+\frac{1}{2})}(ka) \ . \tag{19.115}$$

Les équations (19.113) et (19.115) constituent un système linéaire inhomogène permettant (par exemple) d'exprimer C_{E+} et C_{E-} en fonction de C_{El}. Visiblement, ces constantes sont de la forme $C_{E\pm} = A_\pm(E)\, C_{El}$, où $A_\pm(E)$ est une fonction déterminée de l'énergie. Le déterminant du système est le Wronskien des deux fonctions de Bessel, $J_\nu(z) J'_{-\nu}(z) - J'_\nu(z) J_{-\nu}(z)$ avec $\nu = l + \frac{1}{2}$, égal à $(-1)^{l+1} \frac{2}{\pi k a}$ ([38], § 8.474), et on obtient :

$$C_{E+} = (-1)^{l+1} \frac{\pi}{2} [ka J_{l+\frac{1}{2}}(Ka) J'_{-l-\frac{1}{2}}(ka) - Ka J_{-l-\frac{1}{2}}(ka) J'_{l+\frac{1}{2}}(Ka)]\, C_{El} \ ; \tag{19.116}$$

$$C_{E-} = (-1)^{l+1} \frac{\pi}{2} [Ka J_{l+\frac{1}{2}}(ka) J'_{l+\frac{1}{2}}(Ka) - ka J_{l+\frac{1}{2}}(Ka) J'_{l+\frac{1}{2}}(ka)]\, C_{El} \ . \tag{19.117}$$

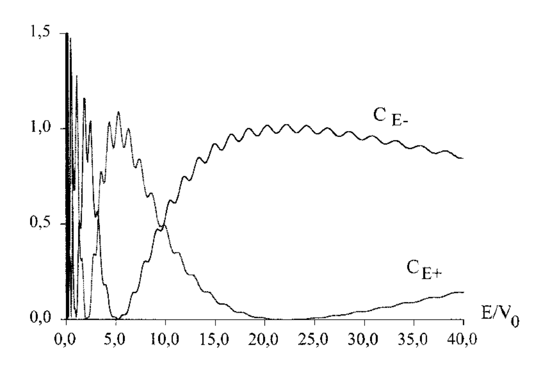


Figure 19.1: Variation générale des coefficients $C_{E\pm}$ en fonction de l'énergie pour $l = 0$ (onde "s") et pour un puits moyennement profond ($k_0 a = 15$).

Ceci achève de déterminer complètement la solution, à un facteur global C_{El} près ; cette indétermination n'en est pas une : pour la description physique de la diffusion par le puits, il faudra tôt ou tard former des paquets d'ondes, de sorte que le choix de la normalisation des fonctions de base est une pure affaire de convenance. Par ailleurs, aucune autre condition n'est encore à prendre en compte, le nombre d'onde k reste quelconque et l'énergie n'est donc pas quantifiée.

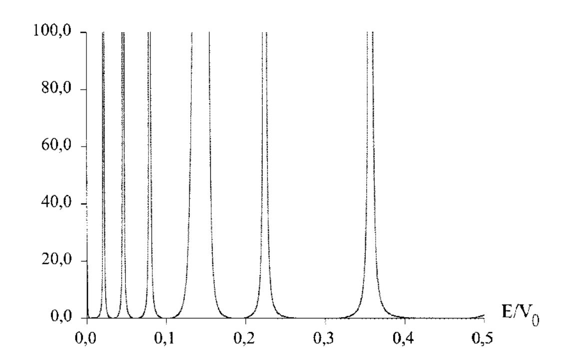


Figure 19.2: Variation du rapport $\left(\frac{C_{E+}}{C_{E-}}\right)^2$ à basse énergie pour $l = 1$ (onde "P") pour un puits profond ($k_0 a = 50$).

D'une façon générale, la variation des coefficients $C_{E\pm}$ en fonction de l'énergie est fort complexe, comme en atteste le tracé de la figure 19.1, et présente une grande

variabilité en fonction du paramètre $k_0 a$, lequel conditionne le nombre d'états liés du puits (voir (19.106)). Pour $k_0 a$ assez grand, les courbes sont fortement structurées et modulées.

On pressent que les coefficients C_{E+} et C_{E-} seront intimement liés à certains coefficients de réflexion et de transmission par le puits carré, plus précisément à la section efficace de diffusion (voir chapitre 26). En particulier, leur variation à basse énergie révèle, tout comme dans le cas unidimensionnel, l'existence de résonances à basse énergie donnant une image, à énergie *positive* des états liés du puits (voir Tome I, chapitre 15, sous-section 15.3.1). Ce phénomène est visible sur la figure 19.2 où est tracé le rapport $(C_{E+}/C_{E-})^2$ pour des petites énergies, $l = 1$ et pour un puits profond ($k_0 a = 50$), et sera discuté plus en détail dans le chapitre 26 (voir en particulier section 26.5 et fig. 26.3, p. 1211). La figure 19.3 donne la variation du même rapport $(C_{E+}/C_{E-})^2$ en fonction de l'énergie dans le cas d'un puits peu profond ($k_0 a = 3$) ; on note les ondulations autour d'une variation moyenne linéaire, provenant des oscillations des fonctions de Bessel.

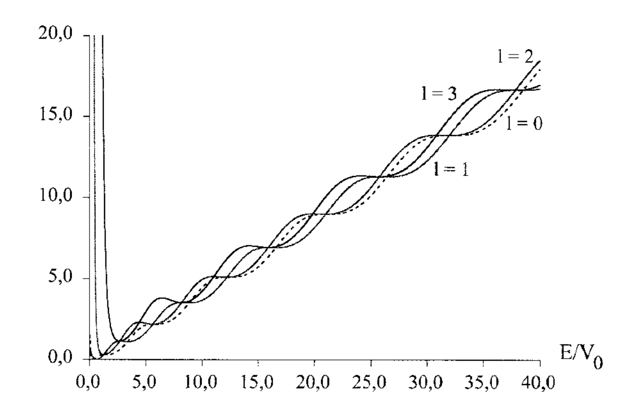


Figure 19.3: Variation du rapport $\left(\frac{C_{E+}}{C_{E-}}\right)^2$ en fonction de l'énergie pour quelques valeurs de l (les courbes sont tracées avec $k_0 a = 3$).

En utilisant la relation $zJ'_\nu(z) = \pm zJ_{\nu\mp 1} \mp \nu J_\nu(z)$, on peut récrire les $C_{E\pm}$ sous une forme où ne figurent plus les dérivées des fonctions de Bessel, ce qui est utile pour l'analyse à haute énergie :

$$C_{E+} = (-1)^l \frac{\pi}{2}[kaJ_{l+\frac{1}{2}}(Ka)J_{-l+\frac{1}{2}}(ka) + KaJ_{-l-\frac{1}{2}}(ka)J_{l-\frac{1}{2}}(Ka)]\,C_{El} \ ; \qquad (19.118)$$

$$C_{E-} = (-1)^l \frac{\pi}{2}[kaJ_{l+\frac{1}{2}}(Ka)J_{l-\frac{1}{2}}(ka) - KaJ_{l+\frac{1}{2}}(ka)J_{l-\frac{1}{2}}(Ka)]\,C_{El} \ . \qquad (19.119)$$

En utilisant les développements asymptotiques des fonctions de Bessel, on trouve tout naturellement, puisque dans ces conditions le puits devient quasi invisible, que C_{E+} tend

vers C_{El}, tandis que C_{E-} tend vers zéro. Le calcul donne précisément :

$$C_{E+} = \left[1 - \mathcal{O}\Big(\frac{V_0}{E}\Big)\right] C_{El} \ , \quad C_{E-} = \frac{(-1)^l}{2} k_0 a \sqrt{\frac{V_0}{E}}\, C_{El} \qquad (E \gg V_0) \ . \tag{19.120}$$

soit :

$$\boxed{\left(\frac{C_{E+}}{C_{E-}}\right)^2 \simeq \frac{4}{(k_0 a)^2}\frac{E}{V_0} = \frac{2\hbar^2 E}{m a^2 V_0^2} \qquad (E \gg V_0)} \tag{19.121}$$

La variation quasi-linéaire à haute énergie est visible sur la figure 19.3.

♦ *Remarques*

1. Si la profondeur du puits V_0 est très grande devant toute autre énergie, et si son rayon a est beaucoup plus petit que toute autre longueur, il est tentant – comme on le fait sans difficulté à une dimension – de modéliser le puits par une fonction de Dirac. Une telle tentative se révèle problématique ([41], [42]) car la réponse dépend du processus de limite. Si on pose :

$$V(r) = A\,\delta(\vec{r}) \quad , \tag{19.122}$$

la constante A a pour dimension [énergie] × [longueur]^D, où D est la dimension de l'espace physique ; compte tenu des seules échelles disponibles, il est licite de poser $A = V_0 a^D$, de sorte que[31] :

$$V(r) = V_0 a^D\, \delta(\vec{r}) \ . \tag{19.124}$$

Formellement, la limite du puits δ s'obtient en faisant $a \to 0$ et $V_0 \to +\infty$; pour que la modélisation garde un sens, ces deux limites ne doivent pas être indépendantes : il faut $V_0 a^D = \mathrm{C^{ste}}$. En conséquence, la combinaison $V_0 a^2$ apparaissant dans (19.106) est proportionnelle à a^{2-D}, montrant que la régularisation est nécessaire dès que $D \geq 2$, ce que confirme l'analyse détaillée [43] .

2. Un autre exemple de champ central intéressant est celui de l'oscillateur harmonique à trois dimensions – qui ne possède que des états liés puisque $V(r \to \infty) \to +\infty$. Ce problème se sépare évidemment en trois oscillateurs indépendants le long de chaque direction cartésienne. Dans le cas où l'oscillateur est isotrope (même constante de raideur suivant les trois directions), la symétrie sphérique donne également lieu à la séparation des variables sphériques[32], fournissant une nouvelle illustration de deux choix différents d'ensembles d'observables qui commutent : en cartésiennes, les éléments de l'ECOC sont les trois énergies de vibration le long des trois axes, dont

[31] L'expression (19.124) est la première qui vient à l'esprit. En fait, on pourrait tout autant poser :

$$V(r) = V_0 a^D f(\lambda)\, \delta(\vec{r}) \tag{19.123}$$

où f est une fonction sans dimension d'un paramètre adimensionné contenant a, par exemple $\lambda = Ka$. L'analyse précise du puits *fini* en dimension $D = 2$ permet de trouver la fonction f et de montrer que, dans ce cas, la modélisation régularisée introduit un puits δ “ramolli” (pour plus de détails, voir [43]).

[32] Voir [25], complément $\mathrm{B_{VII}}$.

la somme est lénergie totale E ; en sphériques, l'ECOC est constitué par l'énergie totale E, $\vec{L}^2$ et L_z. À nouveau, pour une énergie donnée, les deux types d'états propres se déduisent les uns des autres par des combinaisons linéaires exprimant la décomposition des vecteurs propres d'un ECOC sur les états propres de l'autre (voir problème 19.3.10). Tout comme le champ Coulombien, un tel oscillateur possède une dégénérescence *accidentelle* (voir section 19.2.3), phénomène classique quand il existe plusieurs jeux de coordonnées séparables[33].

3. Les quelques exemples abordés montrent bien que les problèmes à trois dimensions sont nettement plus complexes que ceux à une dimension – notamment, le simple puits carré introduit déjà des fonctions spéciales. En outre, tout ce que l'on sait des propriétés générales du mouvement à une dimension ne se généralise pas forcément à trois dimensions. Par exemple, alors que tout état lié à une dimension[34] est non-dégénéré, ce n'est visiblement pas le cas dans $\mathbb{R}^3$ (pour un potentiel à symétrie sphérique, il y a déjà la dégénérescence liée à la symétrie sphérique). Autre exemple : le courant de probabilité est nul dans tout état stationnaire à une dimension ; ce n'est plus vrai à trois dimensions : il peut exister des courants permanents (stationnaires) qui "tournent en rond"[35], ce qui n'est pas possible dans $\mathbb{R}$ pour des raisons purement topologiques. ♦

19.2 Atome d'hydrogène

Pour l'atome d'hydrogène, le potentiel est[36] $V(r) = -\frac{e'^2}{r}$; l'équation radiale (19.38) pour la fonction $u(r)$ est donc :

$$\boxed{H_{\text{eff, Coul}}u(r) \stackrel{\text{déf}}{=} -\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2u}{\mathrm{d}r^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{e'^2}{r}\right]u(r) = Eu(r)} \tag{19.125}$$

où l est l'entier naturel[37] donnant la valeur propre $\hbar^2 l(l+1)$ pour le carré du moment cinétique orbital. Le potentiel effectif :

$$V_{\text{eff, Coul}}(r) \stackrel{\text{déf}}{=} \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{e'^2}{r} \tag{19.126}$$

a deux comportements très tranchés selon que $l = 0$ ou que $l > 0$. Dans le premier cas, il n'y a pas de fait de terme centrifuge et $\lim_{r\to 0} V_{\text{eff, Coul}}(r) = -\infty$; au contraire,

[33]Pour le champ Coulombien, l'équation aux valeurs propres se sépare tant en coordonnées sphériques qu'en coordonnées paraboliques (voir aussi note 73 p. 826).

[34]sur *tout* $\mathbb{R}$; sur un intervalle fini ou semi-infini, certaines propriétés de dégénérescence sont modifiées, notamment la double dégénérescence de tout état non-lié.

[35]Dans un état de l'atome d'hydrogène où L_z n'est pas nul, il y a une *boucle de courant* perpendiculaire à Oz (voir p. 819).

[36]Pour un ion hydrogénoïde de charge nucléaire Z, il suffit de remplacer partout e'^2 par Ze'^2.

[37]Rappelons que toute solution $u(r)$ est paramétrée par l'entier l apparaissant dans $H_{\text{eff, Coul}}$ mais, pour la simplicité des notations, cette dépendance est sous-entendue.

$l > 0$ entraîne $\lim_{r\to 0} V_{\text{eff, Coul}}(r) = +\infty$, la composante centrifuge jouant le rôle d'une barrière répulsive floue, d'autant plus raide que le moment cinétique est grand. Cette distinction qualitative est la raison d'une démarcation nette entre les états $l = 0$ et les états $l > 0$: comme il n'existe pas d'effet répulsif pour les premiers, la densité électronique correspondante $|R_{l=0}(r)|^2$ est non-nulle en $r = 0$, alors que, dans tous les autres cas, $|R_{l>0}(0)|^2 = 0$; on verra que $|R_{l>0}(r)|^2 \propto r^{2l}$ pour r petit (voir (19.182)) : la densité électronique est d'autant plus *écrasée* près de $r = 0$ que le moment cinétique est élevé, c'est-à-dire que la barrière centrifuge est violente. Le potentiel effectif $V_{\text{eff, Coul}}(r)$ est minimum en $r_{\min} = l(l+1)\frac{\hbar^2}{\mu e'^2}$:

$$V_{\text{eff, Coul}}(r) \geq V_{\text{eff, Coul}}(r_{\min}) = -\frac{\mu e'^4}{2l(l+1)\hbar^2} \ ; \qquad (19.127)$$

on sait que ceci constitue une borne inférieure pour toute énergie propre :

$$\boxed{E > -\frac{\mu e'^4}{2l(l+1)\hbar^2}} \qquad (19.128)$$

puisqu'il doit toujours y avoir une région accessible classiquement, c'est-à-dire qu'il doit exister un intervalle (compact) pour r où la différence $E - V_{\text{eff}}(r)$ est positive (voir Tome I, éq. (15.25)).

Lorsque l'énergie E est positive, le mouvement classique à la Kepler (hyperbole si $E > 0$, parabole si $E = 0$) n'est pas borné ; le mouvement quantique correspondant est décrit par une fonction d'onde qui oscille à l'infini et représente un état non-lié. Dans le cas contraire, $E < 0$, la particule classique reste confinée entre deux valeurs r_- et r_+ (ellipse, ou cercle quand $r_- = r_+$) : l'équivalent quantique est un état lié dont la fonction d'onde décroît essentiellement comme une exponentielle quand r tend vers l'infini.

19.2.1 Résolution de l'équation radiale

États liés

Comme toujours, les conditions requises pour la fonction d'onde provoquent la quantification spontanée de l'énergie des états liés. Elles ont été vues à l'œuvre à de multiples reprises dans le Tome I (voir en particulier les chapitres 15 et 16), et à propos des harmoniques sphériques (voir section 18.3) ; dans ce dernier cas, c'est le caractère *monovalué* de la fonction d'onde qui engendre la quantification de la composante L_z du moment cinétique. De même, la condition de normalisabilité de la fonction radiale $R(r)$, et la condition aux limites en $r = 0$, ne seront satisfaites que si l'énergie prend ses valeurs dans un ensemble discret $\{E_n\}_n$, isomorphe à $\mathbb{N}$ pour le potentiel hydrogénoïde. Il sera en outre un peu surprenant de constater, en bout de course, que la formule donnant l'énergie est exactement la même que celle issue de la théorie de Bohr : si on doit s'attendre à ce que les deux expressions de l'énergie tendent l'une vers l'autre dans la limite des grands

nombres quantiques, différant l'une de l'autre par des corrections tendant vers zéro quand $n \to \infty$, rien ne semble devoir assurer d'avance qu'elles soient identiques $\forall n$. C'est pourtant ce que l'on va trouver : les deux expressions de l'énergie coïncident quel que soit n, petit ou grand.

Pour les états liés, l'énergie E doit être négative ; de fait, quand r tend vers l'infini, le terme potentiel est négligeable et il reste :

$$-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2 u}{\mathrm{d}r^2} \simeq E\,u(r) \ . \tag{19.129}$$

Pour que la solution u soit normalisable[38], il est nécessaire qu'elle tende vers zéro à l'infini : si E était positif, u – donc $|u|^2$ – oscillerait sans s'amortir ; au contraire, avec E négatif, (19.129) montre que le comportement est essentiellement exponentiel, conformément à ce qui a été montré dans la sous-section 19.1.2.

Pour trouver les bonnes solutions physiques de cette équation, on utilise la méthode polynômiale comme cela a été fait à propos de l'oscillateur harmonique à une dimension[39]. Comme d'habitude, afin de simplifier le plus tôt possible le problème mathématique sur la base de considérations physiques, posons-nous la question du comportement précis à l'infini des solutions ; ceci permet d'écarter d'emblée celles qui ne sont pas acceptables. Reportons dans (19.125) un comportement du type :

$$u(r) \simeq r^\lambda\,\mathrm{e}^{-kr} \quad (r \to +\infty) \ ; \tag{19.130}$$

ne retenant que les termes dominants[40], on obtient :

$$-\frac{\hbar^2 k^2}{2\mu}\,r^\lambda\,\mathrm{e}^{-kr} \simeq E\,r^\lambda\,\mathrm{e}^{-kr} \ . \tag{19.131}$$

E étant négatif, la forme simplifiée (19.130) est donc solution si $k = \pm\sqrt{-2\mu E/\hbar^2}$; les seules solutions physiques se comportent donc comme suit à l'infini :

$$u(r) \simeq r^\lambda\,\mathrm{e}^{-kr} \ , \qquad k = \sqrt{-2\mu E/\hbar^2} > 0 \qquad (r \to +\infty) \ , \tag{19.132}$$

l'exposant λ étant pour l'instant indéterminé (il peut visiblement être positif ou négatif).

Par ailleurs, on sait d'après (19.60) que le comportement de u à l'origine est :

$$u(r) \simeq r^{l+1} \qquad (r \simeq 0) \ . \tag{19.133}$$

Toutes ces considérations suggèrent des changements de variable et de fonction inconnue. Outre le changement évident $\rho = kr$ introduisant une variable radiale adimensionnée, il

[38]La normalisabilité s'exprime par $\int_0^{+\infty} |u(r)|^2\,\mathrm{d}r < \infty$: une condition *nécessaire* pour cela est $\lim_{r\to+\infty} |u(r)|^2 = 0$. Rappelons toutefois qu'il existe des fonctions qui ne tendent pas vers zéro et qui sont (simplement) intégrables, par exemple $\cos x^2$ (qui conduit à l'intégrale de Fresnel).

[39]voir Tome I, chapitre 16.

[40]Le terme dominant venant de la dérivée seconde est $r^\lambda\,\mathrm{e}^{-kr}$. Les termes en r^{-1} et r^{-2} venant du potentiel effectif dans (19.125) sont sous-dominants.

est en effet commode d'exhiber les comportements aux limites maintenant déterminés en introduisant une nouvelle fonction $w(\rho)$:

$$u(r) = \rho^{l+1}\,\mathrm{e}^{-\rho}\,w(\rho) \ , \qquad \rho = kr \ . \tag{19.134}$$

La fonction $w(\rho)$ peut diverger, mais pas trop vite afin de respecter la condition nécessaire $\lim_{r\to+\infty} u(r) = 0$. En outre, la fonction $w(\rho)$ doit être finie à l'origine puisque le comportement de u, connu d'après (19.60), a été mis en exergue : il faut donc $w(\rho) \sim \mathrm{C}^{\text{ste}}$ quand $\rho \to 0$. La substitution de w à u selon (19.134) dans l'équation complète (19.125) donne :

$$\rho\,\frac{\mathrm{d}^2 w}{\mathrm{d}\rho^2} + 2(l+1-\rho)\,\frac{\mathrm{d}w}{\mathrm{d}\rho} + [\lambda - 2(l+1)]\,w = 0 \tag{19.135}$$

avec :

$$\lambda = \sqrt{\frac{2\mu e'^4}{-\hbar^2 E}} = \frac{2\mu e'^2}{\hbar^2 k} \ . \tag{19.136}$$

L'équation (19.135) est un cas particulier de l'équation de Laplace, dont la forme canonique pour une fonction $L(z)$ est :

$$\boxed{z\,\frac{\mathrm{d}^2 L}{\mathrm{d}z^2} + (\gamma - z)\,\frac{\mathrm{d}L}{\mathrm{d}z} - \alpha\,L = 0} \tag{19.137}$$

où α et γ sont des constantes complexes quelconques. Il suffit en effet de poser :

$$w(\rho) = L(z = 2\rho) \tag{19.138}$$

pour que l'équation (19.135) devienne :

$$z\,\frac{\mathrm{d}^2 L}{\mathrm{d}z^2} + [2(l+1) - z]\,\frac{\mathrm{d}L}{\mathrm{d}z} + (\frac{\lambda}{2} - l - 1)\,L = 0 \ ; \tag{19.139}$$

par comparaison avec (19.135), on a $\alpha = -\frac{\lambda}{2} + l + 1$ et $\gamma = 2l + 2$. L'équation de Laplace (19.137) possède une solution analytique à l'origine[41], qui est une fonction spéciale de première importance, la fonction hypergéométrique dégénérée[42] ([38], § 9.21), notée $\Phi(\alpha,\,\gamma,\,z)$, que l'on retrouvera souvent dans la suite. En tant que fonction analytique dans le voisinage de l'origine, $\Phi(\alpha,\,\gamma,\,z)$ admet un développement[43] en série entière

[41] L'autre solution linéairement indépendante de (19.137) est la fonction $z^{1-\gamma}\Phi(\alpha - \gamma + 1,\, 2 - \gamma,\, z)$, qui se comporte comme $\rho^{-(2l+1)}$ pour $\rho \ll 1$ et n'est donc pas acceptable. L'équation de Laplace se résout commodément en utilisant la méthode de Laplace ([48], Appendice, §a, [4], chapitre 10, § 7, et problèmes 19.3.12 et 19.3.13).

[42] On dit aussi *confluente*, ou fonction de Kummer ([37], § 13.1.2).

[43] En utilisant la relation fonctionnelle $\Gamma(z+1) = z\Gamma(z)$ de la fonction d'Euler de seconde espèce, on voit que le développement de $\Phi(\alpha,\,\gamma,\,z)$ s'écrit aussi :

$$\Phi(\alpha,\,\gamma,\,z) \stackrel{\text{déf}}{=} \sum_{n=0}^{+\infty} \frac{\Gamma(\alpha+n)\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma+n)}\,\frac{z^n}{n!} \ . \tag{19.140}$$

Pour $\Re z > 0$, on peut définir $\Gamma(z)$ comme égale à l'intégrale :

$$\Gamma(z) \stackrel{\text{déf}}{=} \int_0^{+\infty} t^{z-1}\,\mathrm{e}^{-t}\,\mathrm{d}t \qquad (\Re z > 0) \ , \tag{19.141}$$

mais cette fonction peut être prolongée dans $\mathbb{C}$. Elle possède des pôles simples pour $z \in -\mathbb{N}$; par ailleurs, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

de la variable z, qui s'écrit :

$$\Phi(\alpha,\,\gamma,\,z) \stackrel{\text{déf}}{=} 1 + \frac{\alpha}{\gamma}\frac{z}{1!} + \frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{z^2}{2!} + \frac{\alpha(\alpha+1)(\alpha+2)}{\gamma(\gamma+1)(\gamma+2)}\frac{z^3}{3!} + \ldots \ , \tag{19.142}$$

et converge pour tout z fini ; noter que si α est un entier négatif $-q$, Φ se réduit à un polynôme de degré q. Il est facile de vérifier, par substitution directe, que $\Phi(\alpha,\,\gamma,\,z)$ satisfait bien l'équation (19.137).

Posons donc pour $w(\rho)$ un développement en série entière ne contenant que des termes de puissance positive ou nulle :

$$w(\rho) = \sum_{p=0}^{+\infty} c_p \rho^p \ , \tag{19.143}$$

avec $c_0 \neq 0$ puisque r^{l+1} étant précisément le comportement de la fonction u près de $r = 0$, la série doit se comporter comme une constante dans ce voisinage. Le report dans l'équation (19.135) fournit la récurrence suivante[44] :

$$(p+1)(p+2l+2)\,c_{p+1} + [\lambda - 2(l+1+p)]\,c_p = 0 \ , \quad p \in \mathbb{N} \ , \quad c_0 \neq 0 \ . \tag{19.144}$$

Ceci permet bien d'exprimer tous les coefficients en fonction du premier c_0 qui, sans autre condition supplémentaire à utiliser en temps utile, est arbitraire puisque l'équation différentielle est homogène. La relation (19.144) définit donc complètement (à un facteur près) l'ensemble des solutions mathématiques qui sont *régulières* à l'origine. Cette régularité (analycité) n'est toutefois pas suffisante pour élire toutes ces solutions comme physiquement acceptables : on va en effet montrer que la série (19.143) se comporte comme $\mathrm{e}^{+2\rho}$ à l'infini ; elle ne peut donc conduire à une solution convenable que si tous ses coefficients sont nuls à partir d'un certain rang, auquel cas la série se réduit alors à un *polynôme* ; quel que soit le degré de ce polynôme, l'exponentielle décroissante dans (19.134) finit par l'emporter et u, tendant vers zéro à l'infini, possède le bon comportement nécessaire ; de surcroît, décroissant exponentiellement, cette fonction est visiblement de module carré sommable : la condition nécessaire se révèle ainsi être aussi suffisante.

La divergence exponentielle de la série (19.143) peut être mise en évidence de plusieurs façons. Donnons en premier un argument très intuitif, facile à formaliser d'ailleurs pour en faire apparaître le bien-fondé. À l et E fixés, pour p grand, la récurrence (19.144) s'écrit à peu près :

$$p^2\,c_{p+1} - 2p\,c_p \simeq 0 \quad \Longleftrightarrow \quad \frac{c_{p+1}}{c_p} \simeq \frac{2}{p} \ . \tag{19.145}$$

[44]On peut étendre les valeurs de p aux valeurs entières négatives à condition de garder en tête que $c_{-1} = 0$. Par ailleurs, la présence du facteur $(p+1)$ pour c_{p+1} assure le découplage des $c_{p\geq 0}$ et des éventuels $c_{p<0}$; ce fait traduit l'existence de deux solutions linéairement indépendantes – comme pour toute équation du second ordre. L'une est régulière à l'origine, c'est celle que l'on retient ici ; l'autre est divergente en $\rho = 0$, mise à l'écart par le choix pragmatique (19.134) inspiré par les conditions physiques imposées aux solutions à retenir exclusivement.

On connaît une série dont le rapport de deux termes consécutifs est $2/p$, c'est $e^{2\rho}$:

$$e^{2\rho} = \sum_{p=0}^{+\infty} \frac{1}{p!}\,(2\rho)^p \equiv \sum_{p=0}^{+\infty} \gamma_p\,\rho^p, \qquad \frac{\gamma_{p+1}}{\gamma_p} = \frac{2}{p}\ . \tag{19.146}$$

Pour ρ tendant vers l'infini, ce sont les grandes valeurs de p qui comptent ; il en résulte que, tant que w est donnée par la série entière, w se comporte à peu près comme $e^{2\rho}$ quand $\rho \to +\infty$. Une autre façon d'établir ce fait consiste à poser que pour ρ grand, on a $w \sim \rho^q\,e^{\alpha\rho}$. Le report dans (19.135) donne, après simplification par $\rho^{q-1}e^{\alpha\rho}$:

$$q(q-1) + 2q\alpha\rho + \alpha^2\rho^2 + 2(l+1-\rho)(q+\alpha\rho) + (\lambda - 2l - 2)\rho \sim 0\ . \tag{19.147}$$

Seuls les termes les plus divergents doivent être considérés[45], soit le terme en ρ^2 ; son coefficient, égal à $(\alpha^2 - 2\alpha)$, doit être nul, ce qui fournit bien la seule valeur acceptable[46] $\alpha = 2$, confirmant l'analyse fondée sur la comparaison des séries.

Donnons un dernier argument pour établir définitivement le fait crucial que la série ne doit pas être conservée – crucial puisque c'est ce qui produit spontanément la *quantification* de l'énergie des états liés. La comparaison des équations (19.135) et (19.137) montre que :

$$\frac{1}{c_0}w(\rho) = \Phi(l+1-\frac{\lambda}{2},\, 2l+2,\, 2\rho)\ . \tag{19.148}$$

On connaît par ailleurs le développement (asymptotique) à l'infini de $\Phi(\alpha,\,\gamma,\,z)$ ([37], § 13.5.1) :

$$\Phi(l+1-\frac{\lambda}{2},\, 2l+2,\, 2\rho) \sim \frac{\Gamma(l+1+\frac{\lambda}{2})}{\Gamma(l+1-\frac{\lambda}{2})}\,(2\rho)^{-(l+1+\frac{\lambda}{2})}\,e^{2\rho} \qquad (\rho \gg 1)\ , \tag{19.149}$$

d'où l'on déduit précisément que, pour $r \gg k^{-1}$, $u(r)$ se comporte comme :

$$u(r) \sim (kr)^{-\frac{\lambda}{2}}\,e^{-kr}\,e^{+2kr} = (kr)^{-\frac{\lambda}{2}}\,e^{+kr} \qquad (r \gg k^{-1}) \tag{19.150}$$

Ainsi, d'une façon ou d'une autre, on constate que si l'on garde la série (19.143) telle quelle, la fonction $u(r)$ diverge exponentiellement quand r tend vers l'infini. La seule et unique façon de soigner cette divergence inacceptable physiquement est d'annuler tous les coefficients c_p à partir d'un certain rang : alors la série se réduit à un certain *polynôme* ; aussi grand que soit le degré de celui-ci, l'exponentielle finit toujours par l'emporter, entraînant $\lim_{r\to+\infty} u(r) = 0$ en conformité avec (19.41). Pour qu'il en soit ainsi, la seule possibilité est qu'il existe un entier p_0 tel que $\lambda - 2l - 2 - 2p_0 = 0$, donnant $c_{p_0+1} = c_{p_0+2} = \ldots = 0$. Dans ces conditions, et selon (19.134), le comportement de $u(r)$ à l'infini est :

$$u(r) \simeq (kr)^{l+1+p_0}\,e^{-kr} \qquad (r \gg k^{-1})\ . \tag{19.151}$$

[45]Les autres devraient être comparés avec des termes omis en posant $w \sim \rho^q\,e^{\alpha\rho}$.

[46]L'équation $\alpha^2 - 2\alpha = 0$ a aussi la solution $\alpha = 0$, auquel cas la fonction $w(\rho)$ se comporte à l'infini comme un polynôme de degré q, indéterminé à ce stade. Compte tenu de la conclusion qui va suivre, retenir cette solution n'enrichit donc pas l'ensemble des solutions acceptables.

Cela étant acquis, les seules bonnes solutions sont donc celles pour lesquelles :

$$\lambda - 2l - 2 = 2p_0 \ , \qquad p_0 \in \mathbb{N}. \tag{19.152}$$

p_0 et l étant des entiers positifs ou nuls, les valeurs possibles de λ sont les entiers pairs strictement positifs[47], ce qui conduit à poser :

$$\lambda = 2n \ , \qquad n \in \mathbb{N}^*. \tag{19.153}$$

Selon (19.136), les seules valeurs de k à retenir sont les valeurs *discrètes* :

$$\boxed{k \in \{k_n\}_n \ : \quad k_n = \frac{\mu e'^2}{n\hbar^2} \equiv \frac{1}{n\tilde{a}_0} \qquad (n \in \mathbb{N}^*)} \tag{19.154}$$

où on a défini :

$$\boxed{\tilde{a}_0 \stackrel{\text{déf}}{=} \frac{\hbar^2}{\mu e'^2}} \tag{19.155}$$

Autrement dit, et toujours selon (19.136), on obtient le résultat essentiel attestant de la quantification de l'énergie :

$$\boxed{E \in \{E_n\}_n \ : \quad E_n = -\frac{\mu e'^4}{2n^2\hbar^2} \equiv -\frac{e'^2}{2n^2\tilde{a}_0} \qquad (n \in \mathbb{N}^*)} \tag{19.156}$$

Ainsi, une fois de plus, la condition de normalisabilité des fonctions d'onde des états liés vient faire le tri parmi les solutions mathématiques de l'équation aux valeurs propres. Tout comme la condition de Bohr - Wilson - Sommerfeld $\oint p\,\mathrm{d}q =$ entier $\times\, h$ le faisait parmi toutes les solutions classiques, la condition de normalisabilité passe au peigne fin les solutions mathématiques déjà triées pour en extraire un ensemble dénombrable, ici infini, achevant la démonstration de la quantification de l'énergie des états liés.

Les fonctions radiales $R(r) = r^{-1}u(r)$, qui dépendent évidemment de l'entier l figurant explicitement dans l'équation différentielle (19.125), seront également repérées par l'indice n associé à la valeur E_n de l'énergie ; on notera ainsi[48] $R_{nl}(r)$ la fonction radiale qui, multipliée par Y_{lm}, constitue un état propre lié normalisé de $(H, \vec{L}^2, L_z)$ pour l'atome d'hydrogène[49] :

$$\boxed{\psi_{nlm}(r,\,\theta,\,\phi) = R_{nl}(r)Y_{lm}(\theta,\phi)} \tag{19.157}$$

avec :

$$\boxed{R_{nl}(r) = C_{nl}\,(2k_n r)^l\,\mathrm{e}^{-k_n r}\Phi(l+1-n, 2l+2, 2k_n r)} \tag{19.158}$$

[47] Avec $\alpha = n - (l+1)$, on voit bien sur la série (19.142) et compte tenu de l'identification (19.148), que Φ, donc w, se réduit à un polynôme de degré l.

[48] Une autre notation, plus naturelle, serait $R_{E_n l}(r)$; on s'en tiendra ici à la notation consacrée par l'usage.

[49] À nouveau, une notation plus naturelle est $\psi_{k_n lm}(r,\,\theta,\,\phi)$.

et $n = \frac{1}{k_n \tilde{a}_0}$. Comme $l + 1 - n$ ($\equiv \alpha$) est ici un entier négatif, la fonction Φ est bien un certain polynôme de degré $n - l - 1$, appelé *polynôme associé de Laguerre*[50], dont tous les zéros sont réels et faisant l'objet de l'aparté technique p. 820. Par ailleurs, les harmoniques sphériques étant normalisées, la fonction $R_{nl}(r)$ est normalisée à part, satisfaisant :

$$\int_0^{+\infty} [R_{nl}(r)]^2 \, r^2 \, \mathrm{d}r = 1 \ , \tag{19.159}$$

ce qui fixe le facteur C_{nl} (voir (19.182)).

L'expression (19.156) de l'énergie est strictement identique à celle de la théorie de Bohr, dont on sait qu'elle reproduit[51] la formule de Balmer pour les termes spectraux de l'hydrogène, compte tenu de la correction de masse réduite. Comme annoncé, cette coïncidence *quel que soit n*, petit ou grand, a quelque chose de surprenant[52]. Par ailleurs, se souvenir que les fonctions $R_{nl}(r)$ et $R_{nl'}(r)$ (ou u_{nl} et $u_{nl'}$) n'ont nullement à être orthogonales (et d'ailleurs ne le sont pas) ; en revanche, on a (une fois la normalisation effectuée) :

$$\int_0^{+\infty} u_{nl}(r) u_{n'l}(r) \, \mathrm{d}r = \int_0^{+\infty} R_{nl}(r) R_{n'l}(r) \, r^2 \mathrm{d}r = \delta_{nn'} \ . \tag{19.160}$$

États non-liés

Par opposition avec le cas lié, les solutions de (19.125) d'énergie *E positive* ont un comportement oscillant à l'infini, comme le montre immédiatement la forme asymptotique

[50] Edmond Nicolas Laguerre, (1834-1886), est surtout connu pour l'introduction des polynômes qui portent son nom. Il a aussi inventé un algorithme très rapide pour trouver le zéro d'un polynôme (*Méthode de Laguerre*). On lui doit aussi un théorème exprimant dans le cas réel celui de Gauss - Lucas (*"Les zéros de la dérivée $P'(x)$ d'un polynôme sont situés dans l'enveloppe convexe de l'ensemble des racines de $P(x)$"*).

[51] On se souvient que Bohr a construit sa théorie pour qu'il en soit ainsi (voir Tome I, chapitre 6).

[52] Cette coïncidence a réellement quelque chose d'énigmatique. L'auteur avoue son ignorance en n'ayant pas connaissance d'une discussion dans la littérature sur ce point.

Par ailleurs, on peut dès à présent noter une propriété remarquable de l'expression de l'énergie (19.156) : elle ne dépend pas du nombre quantique l associé au module du moment cinétique orbital. Cette propriété est liée à la forme précise du potentiel Coulombien pur et donne lieu à ce qui est appelé souvent – mais incorrectement car elle peut être une source de contre-sens – dégénérescence *accidentelle*. En réalité, cette dégénérescence résulte d'une symétrie remarquable et subtile du potentiel en $1/r$, qui donne lieu à la conservation dynamique du vecteur de Lenz - Runge (voir sous-section 19.2.3, et note 73 p. 826). Cette symétrie, dans l'espace des impulsions (on peut la mettre en évidence en passant en représentation-p), a été remarquée par Pauli [44] puis exploitée par Fock en 1935 [45]. Plus récemment elle a été étudiée notamment par McIntosh [46] et par Bander et Itzykson [47]. En outre, cette dégénérescence particulière est liée au fait qu'il existe un autre jeu de coordonnées (les coordonnées paraboliques) pour lequel les variables spatiales se séparent aussi – ces coordonnées sont d'ailleurs particulièrement bien adaptées au traitement exact (non-perturbatif) de l'effet Stark.

Tout écart à la loi en $1/r$ – par exemple un terme en $1/r^2$ déstabilisant (parfois utilisé pour représenter un écrantage à longue portée) –, supprime cette propriété et l'énergie devient alors une fonction explicite de n *et* de l (voir problème 19.3.15).

(19.129) : ces solutions correspondent à des états propres non-liés. En posant :

$$\kappa = \frac{\sqrt{2\mu E}}{\hbar} \ , \qquad \Lambda = \sqrt{\frac{2\mu e'^4}{\hbar^2 E}} \ , \tag{19.161}$$

et en définissant la fonction $W(\rho)$ par :

$$u(r) = \rho^{l+1} \mathrm{e}^{-\mathrm{i}\rho}\, W(\rho) \ , \qquad \rho = \kappa r \ , \tag{19.162}$$

on trouve que $W(\rho)$ satisfait l'équation (comparer avec (19.135)) :

$$\rho\, \frac{\mathrm{d}^2 W}{\mathrm{d}\rho^2} + 2(l+1-\mathrm{i}\rho)\, \frac{\mathrm{d}W}{\mathrm{d}\rho} + [\Lambda - 2\mathrm{i}(l+1)]\, W = 0 \ . \tag{19.163}$$

Il s'agit encore évidemment d'un avatar de l'équation de Laplace dont la solution analytique[53] en $\rho = 0$, admet un développement en série entière $\sum_{p\in\mathbb{N}} c_p \rho^p$; les coefficients satisfont une relation de récurrence analogue à (19.144), qui s'écrit précisément :

$$(p+1)(p+2l+2)\, c_{p+1} + [\Lambda - 2\mathrm{i}(l+1+p)]\, c_p = 0 \ , \quad p \in \mathbb{N} \ , \quad c_0 \neq 0 \ . \tag{19.164}$$

Posant $\Lambda = 2\nu$ (par analogie avec la *nécessité* $\lambda = 2n$ pour les états liés), on voit que cette solution n'est autre que la fonction $\Phi(l+1+\mathrm{i}\nu,\, 2l+2,\, 2\mathrm{i}\rho)$ définie en (19.142). Par opposition avec le cas des états liés, aucune raison pertinente n'existe conduisant à "briser" la série, la fonction propre n'étant pas tenue de s'annuler à l'infini. Aucun tri n'étant à faire, toutes les valeurs de κ sont acceptables (comparer avec (19.154)) ; selon (19.162), les solutions physiques $u(r)$ sont de la forme (à un facteur près) :

$$u(r) = (\kappa r)^{l+1} \mathrm{e}^{-\mathrm{i}\kappa r}\, \Phi(l+1+\mathrm{i}\nu,\, 2l+2,\, 2\mathrm{i}\kappa r) \qquad (\kappa \in \mathbb{R}_+) \ . \tag{19.165}$$

D'après la discussion p. 786, on sait que, même pour les états non-liés, l'équation aux valeurs propres pour la fonction $u(r)$ ne possède pas de dégénérescence et que, par conséquent, la fonction radiale peut toujours être prise réelle. S'il ne saute pas aux yeux que la fonction écrite en (19.165) est réelle, c'est bien pourtant le cas en vertu de la transformation de Kummer[54] $\Phi(\alpha, \gamma, z) = \mathrm{e}^z \Phi(\gamma - \alpha, \gamma, -z)$, qui donne ici :

$$\begin{aligned} u^*(r) = (\kappa r)^{l+1} \mathrm{e}^{+\mathrm{i}\kappa r}\, \Phi(l+1-\mathrm{i}\nu,\, 2l+2,\, -2\mathrm{i}\kappa r) = \\ (\kappa r)^{l+1} \mathrm{e}^{+\mathrm{i}\kappa r} \mathrm{e}^{-2\mathrm{i}\kappa r}\, \Phi(2l+2-(l+1-\mathrm{i}\nu),\, 2l+2,\, 2\mathrm{i}\kappa r) \equiv u(r) \ . \end{aligned} \tag{19.166}$$

En définitive, la solution la plus générale d'énergie positive $E = \frac{\hbar^2\kappa^2}{2\mu}$ est :

$$\boxed{\psi_{Elm}(r,\, \theta,\, \phi) = R_{El}(r) Y_{lm}(\theta,\, \phi)} \tag{19.167}$$

avec :

$$\boxed{R_{El}(r) = C_{El}\, (2\kappa r)^l\, \mathrm{e}^{-\mathrm{i}\kappa r} \Phi(l+1+\mathrm{i}\nu, 2l+2, 2\mathrm{i}\kappa r)} \tag{19.168}$$

[53] Tout comme dans le cas $E < 0$, il existe une autre solution linéairement indépendante ([38], § 9.216) ; singulière car se comportant comme $\rho^{-(2l+1)}$ en $\rho = 0$, elle est non recevable physiquement.

[54] Voir [37], § 13.1.27 et problème 19.3.12.

où $\nu = \frac{1}{\kappa \tilde{a}_0}$.

À la réflexion et en comparant les deux expressions (19.158) et (19.168), on voit finalement que la solution à énergie positive $R_{El}(r)$ s'obtient par simple prolongement analytique du nombre d'onde $k \stackrel{\text{déf}}{=} \hbar^{-1}\sqrt{-2\mu E}$ des valeurs négatives de E vers les valeurs positives, le paramètre λ abandonnant au passage l'obligation d'être discret. Raisonnant avec le "nombre d'onde" complexe $\tilde{k} \stackrel{\text{déf}}{=} -ik$, on voit que les états liés correspondent à des points d'affixes *discrètes* $-\frac{\mathrm{i}}{n}\tilde{a}_0^{-1}$ ($n \in \mathbb{N}^*$) qui se densifient vers l'origine (point d'accumulation) quand n augmente, c'est-à-dire quand l'énergie E_n augmente. À partir de $E = 0$, $\tilde{k}$ bifurque vers le demi-axe réel positif, et devient une variable *continue.*

Ainsi, l'analyse des états non-liés n'introduit aucune condition de normalisabilité, et donc aucune quantification de l'énergie. Ces états servent de base pour la description en termes de *paquets d'ondes* de la diffusion par le potentiel de Coulomb de particules émises par une source située à l'infini, tout comme dans le cas unidimensionnel, les ondes planes $\mathrm{e}^{\pm ikx}$ (non normalisables) servent de fonctions de base, autorisant l'écriture d'un paquet d'ondes normalisable sous la forme d'une intégrale de Fourier. Ici, un paquet d'ondes pourra toujours s'écrire en combinaison linéaire des états propres, liés et non-liés :

$$\Psi(\vec{r}, t) = \sum_{lm} Y_{lm}(\theta, \phi)\Big(\sum_{n\in\mathbb{N}^*} c_{nlm}\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}E_n t}R_{nl}(r) + \int_0^{+\infty} c_{Elm}\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Et}R_{El}(r)\,\mathrm{d}E\Big) \ , \quad (19.169)$$

les coefficients c_{nlm} et c_{Elm} étant en principe calculables une fois donné l'état initial $\Psi(\vec{r},\, t = 0) \equiv \langle\vec{r}\,|\Psi(0)\rangle$ en effectuant tous les produits scalaires $\langle R_{nl}Y_{lm}|\Psi(0)\rangle$ et $\langle R_{El}Y_{lm}|\Psi(0)\rangle$. Les conditions aux limites à satisfaire, qui prendront en compte notamment la valeur du courant de particules émis par la source et la localisation de celle-ci, ne feront que généraliser ce qui a été vu à une dimension d'espace : dans ce dernier cas, on sait bien que c'est la position de la source et ses caractéristiques[55], qui permettent de déterminer toutes les constantes d'intégration apparaissant dans la résolution de l'équation aux valeurs propres[56]. Il est clair que pour ce calage, le facteur de proportionnalité des fonctions propres non-liées est sans importance, le choix (19.193) n'étant fait (usuellement) que pour la commodité technique. C'est bien l'absence pertinente d'une condition de normalisabilité pour les états propres qui laisse l'énergie des états non-liés à l'écart de la quantification, lui permettant de prendre *toute* valeur E entre 0 et $+\infty$.

Comme expliqué plus haut, l'équation pour la fonction $u(r)$ est une équation aux valeurs propres formellement identique à celle d'un mouvement confiné sur $\mathbb{R}_+$ par une barrière totalement réfléchissante située en $x = 0$, et soumis au potentiel effectif $V_{\text{eff}}(x)$. Avec la condition aux limites de fonctions nulles en $x = 0$, H_{eff} est hermitique et l'ensemble des fonctions propres forme une base complète. Il existe donc une relation de fermeture traduisant cette propriété ; celle-ci contient une sommation sur les états

[55] Par exemple le courant qu'elle émet, voir Tome I, chapitre 15.

[56] Ainsi, pour un potentiel localisé $V(x)$ – nul en-dehors d'un intervalle fini $[a, b]$ –, la solution pour $x > b$ est de la forme $A\mathrm{e}^{ikx} + B\mathrm{e}^{-ikx}$, $k > 0$. Si la source est localisée en $x = -\infty$, la constante B est en fait nulle. Le module de A sera finalement trouvé, *via* les coefficients de réflexion et de transmission, par calage sur le courant de la source $j_{\text{s}} = (\hbar k/m)|A|^2$.

liés *et* sur les états non-liés. Pour une valeur de l donnée, H_{eff} est aussi hermitique, et la relation de fermeture s'écrit :

$$\sum_{n\in\mathbb{N}^*} u_{nl}(r)u_{nl}(r') + \int_0^{+\infty} u_{El}(r)u_{El}(r')\,\mathrm{d}E = \delta(r-r') \ . \tag{19.170}$$

soit :

$$\boxed{\sum_{n\in\mathbb{N}^*} R_{nl}(r)R_{nl}(r') + \int_0^{+\infty} R_{El}(r)R_{El}(r')\,\mathrm{d}E = rr'\delta(r-r')} \tag{19.171}$$

19.2.2 Analyse des états propres

En définitive, le spectre du Hamiltonien coulombien pur se compose donc de deux parties bien distinctes :

1. une partie discrète bornée inférieurement[57] du côté négatif par $E_{n=1}$, constituée de niveaux de plus en plus denses quand n augmente, en nombre infini dénombrable, avec un point d'accumulation en $E=0$. Cette propriété est caractéristique ([34], p. 354) des potentiels qui tendent vers zéro moins vite que r^{-2} ($r^2V(r)\to\infty$ si $r\to+\infty$) ; pour un potentiel $r^2V(r)\to 0$ si $r\to\infty$, allant plus rapidement vers zéro que r^{-2}, le nombre d'états liés est fini (et parfois nul d'ailleurs si le potentiel n'est pas assez attracteur et/ou pas assez étendu dans l'espace) ; le puits carré sphérique appartient à cette catégorie, comme on l'a vu plus haut.
2. une partie continue (dense) s'étendant de 0 à $+\infty$, correspondant aux états non-liés (états stationnaires de diffusion).

L'énergie de l'état fondamental est $E_{n=1}\simeq -13,6$ eV. Par ailleurs, le rayon de la première orbite de Bohr (dans l'approximation du noyau infiniment massif) est[58] :

$$\boxed{a_0 \stackrel{\text{déf}}{=} \frac{\hbar^2}{m_{\text{e}}e'^2} \simeq 0,53\,\text{Å}} \tag{19.172}$$

[57]Le fait que l'énergie E est bornée inférieurement est déjà en soi un succès majeur de la théorie. En effet, un argument simple de stabilité maximale, développé dans un cadre strictement classique s'appuyant sur le théorème du Viriel montre que l'état d'énergie la plus basse est celui où l'électron est … sur le noyau ! Comme alors l'énergie est infinie négative, c'est bien vrai que l'on ne peut pas faire "mieux". L'atome classique n'a vraiment aucune chance d'exister : outre cette instabilité purement mécanique, il est aussi foncièrement instable électrodynamiquement parlant (l'électron, particule accélérée, rayonne et se précipite sur le noyau, voir Tome I, section 1.5 et problème 1.6.9).

[58]m_{e} désigne la masse de l'électron, $m_{\text{e}}\simeq 9\times 10^{-31}$ kg.
Par ailleurs, $\tilde{a}_0 = \frac{m_{\text{e}}}{\mu}\,a_0 \simeq a_0$.

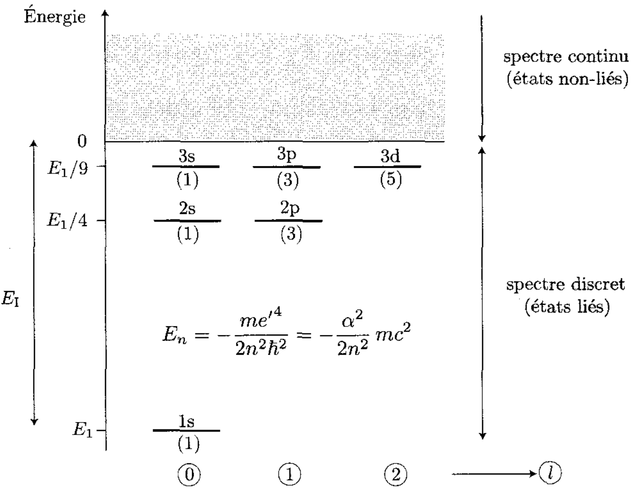


Figure 19.4: Spectre d'énergie de l'atome d'hydrogène et notation spectroscopique. Les chiffres entre parenthèses indiquent la dégénérescence de chaque état. E_I est l'énergie d'ionisation à partir du fondamental, et vaut environ 13, 6 eV.

Une autre expression de l'énergie est souvent utile :

$$\boxed{E_n = -\frac{\alpha^2}{2n^2}\,\mu c^2} \qquad (19.173)$$

où α est la constante de structure fine[59], l'*alpha* des constantes de la Physique :

$$\boxed{\alpha \overset{\text{déf}}{=} \frac{e'^2}{\hbar c} = \frac{1}{137,007\ldots}} \qquad (19.174)$$

Par ailleurs, comme l'entier p_0, rang où la série (19.143) est tronquée, est positif ou nul, les seules valeurs possibles de l pour n fixé satisfont[60] (voir (19.152) et (19.153)) :

$$2n - 2l - 2 \geq 0 \quad \Longleftrightarrow \quad l \leq n-1 \ . \qquad (19.175)$$

On sait déjà que l est positif ou nul ; au total, à n fixé, c'est-à-dire à énergie fixée, les seules valeurs possibles de l sont :

$$\boxed{l = 0,\, 1,\, 2,\, \ldots\,,\, n-1} \qquad (19.176)$$

Physiquement, l'existence d'une borne supérieure pour l se comprend aisément. On sait (Tome I, chapitre 15, éq. (15.24)) qu'il doit toujours exister une région accessible classiquement, avec pour conséquence que toute valeur propre de l'énergie est (strictement)

[59] α contient le produit $\hbar c$: dans toute expression où α figure, l'*ordre* des limites classique ($\hbar \to 0$) et non-relativiste ($c \to +\infty$) est crucial.

[60] Les états $l = 0$ correspondent aux orbites circulaires de la théorie de Bohr.

supérieure au minimum de potentiel (voir (19.128)) ; en vertu de (19.156) et (19.127), il faut donc $l(l+1) < n$, soit $l < n$, ou encore $l \leq n-1$ puisqu'il s'agit d'entiers.

Il apparaît ainsi que l'énergie des états liés de l'atome d'hydrogène ne dépend pas de l ; comme à l fixé il y a $2l+1$ valeurs possibles de m, la dégénérescence totale[61] est :

$$g_n = \sum_{l=0}^{n-1} (2l+1) = 2\,\frac{(n-1)n}{2} + n = n^2 \ . \qquad (19.177)$$

L'indépendance de l'énergie par rapport à l a été brièvement mentionnée plus haut et n'était pas prévisible au vu des seuls arguments de symétrie de rotation ; c'est pourquoi on la qualifie parfois de dégénérescence "accidentelle". Les différents états de l'électron dans l'atome d'hydrogène sont donc entièrement caractérisés par la donnée des trois nombres quantiques n, l et m, tant que l'on maintient le spin à l'écart. Pour des raisons historiques, tenant à la classification "zoologique" des états construite en une période où même la théorie de Bohr n'était pas faite, l'usage est de représenter ce triplet de nombres, ou plus précisément les deux premiers par un symbole appelé *notation spectroscopique*. On le construit en mettant en tête la valeur de n, puis on accole une lettre en correspondance avec la valeur de l :

$$l = 0 \leftrightarrow \mathrm{s} \ , \qquad l = 1 \leftrightarrow \mathrm{p} \ , \qquad l = 2 \leftrightarrow \mathrm{d} \ , \qquad l = 3 \leftrightarrow \mathrm{f} \ , \ \ldots \qquad (19.178)$$

s est pour "sharp", p pour " pure", d pour " diffuse", etc., autant de qualificatifs attribués aux raies spectrales à l'aube de la spectroscopie atomique. Par ailleurs, la notion historique de *couche atomique* (Tome I, p. 234) se réfère à la valeur de n, parfois appelé *nombre quantique principal* puisqu'il donne la valeur de l'énergie (pour l'hydrogène, ou un atome hydrogénoïde). La notation spectroscopique est ainsi : 1s (couche K) ; 2s, 2p (couche L) ; 3s, 3p, 3d (couche M), etc.

En définitive, les fonctions radiales des états liés sont les $R_{nl}(r)$:

$$R_{nl}(r) = C_{nl}\,(2k_n r)^l\, L_{n-l-1}^{2l+1}(2k_n r)\,\mathrm{e}^{-k_n r} \ , \qquad (19.179)$$

où C_{nl} est la constante de normalisation radiale déduite de :

$$\int_0^{+\infty} r^2\, R_{nl}^2(r)\,\mathrm{d}r = 1 \ , \qquad (19.180)$$

sachant que :

$$\int_0^{+\infty} x^q \mathrm{e}^{-x}\,[L_p^q(x)]^2\,\mathrm{d}x = \frac{[(p+q)!]^3}{p!} \ ; \qquad (19.181)$$

(la partie angulaire est normalisée à part). L'expression normalisée explicite des R_{nl} est (voir problème 19.3.12) :

$$\boxed{R_{nl}(r) = \left[4k_n^3\frac{(n-l-1)!}{n[(n+l)!]^3}\right]^{1/2}(2k_n r)^l\, L_{n-l-1}^{2l+1}(2k_n r)\,\mathrm{e}^{-\,k_n r}} \qquad (19.182)$$

[61] Compte non-tenu du spin ; la prise en compte du spin $S = 1/2$ de l'électron ajoute un facteur 2 : la dégénérescence totale est alors égale à $2n^2$, d'où les nombres d'occupation maximaux des couches atomiques K, L, M,... (2, 8, 18, ...).

Les k_n sont les nombres d'onde (19.154), que l'on peut récrire en fonction de a_0 :

$$\boxed{k_n = \frac{1}{1+\frac{m_e}{M_n}} \frac{1}{na_0} \simeq \frac{1}{na_0}} \tag{19.183}$$

M_n désignant la masse du noyau. L_{n-l-1}^{2l+1} est un polynôme de degré $n-l-1$ dont tous les zéros sont simples et réels, en conformité avec le Théorème d'oscillation ([48], § 21 et Tome I, p. 528). Le facteur exponentiel $e^{-k_n r}$ a une portée $k_n^{-1} \simeq na_0$, indiquant que plus l'état est excité, plus la densité de probabilité est diffuse ; toutefois, la variation des L_{n-l-1}^{2l+1} complique la situation, bien que toutes les fonctions de même nombre quantique principal n ont la même échelle spatiale $n\tilde{a}_0$; en tout cas, la fonction R_{nl} s'annule $n-l-1$ fois en dehors de l'origine. La discussion suivante permet de préciser un peu les choses en la matière (voir aussi la question 3 dans le problème 19.3.11).

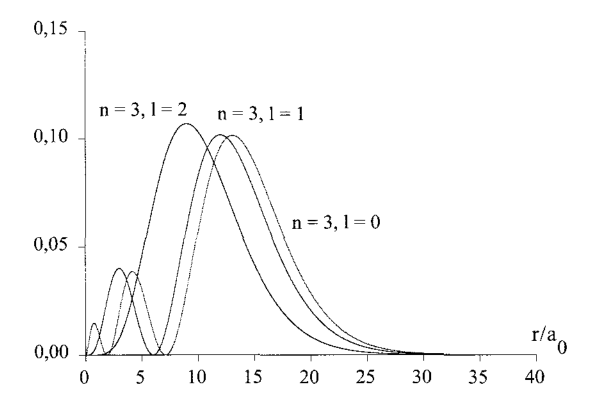


Figure 19.5: Densité radiale $\rho_{nl}(r) \stackrel{\text{déf}}{=} r^2[R_{nl}(r)]^2$ pour l'état lié $n=3$, $l=0, 1, 2$; noter que R_{3l} s'annule $2-l$ fois en dehors de l'origine. En ordonnée est porté le produit $a_0^3\rho_{nl}$.

Le cas où l prend sa valeur maximale $n-1$ est simple à analyser puisqu'alors le polynôme associé de Laguerre est de degré zéro, c'est donc une constante, égale à $(2n-1)!$; dans ces conditions, la fonction radiale est :

$$R_{nl=n-1}(r) = \frac{2a_0^{-3/2}}{n^2\sqrt{(2n-1)!}}\,(2k_n r)^{n-1}\, e^{-k_n r} \ . \tag{19.184}$$

Pour ces états $l = n-1$, la densité de probabilité radiale $\rho_{nl}(r) \stackrel{\text{déf}}{=} r^2[R_{nl}(r)]^2$ est proportionnelle à $r^{2n}\, e^{-2k_n r}$; $\rho_{nl}(r)$ présente un maximum pour $r = r_{\text{max}}$ donné par :

$$r_{\text{max}} = \frac{n}{k_n} = \frac{m_e}{\mu}\, n^2 a_0 \simeq n^2 a_0 \ . \tag{19.185}$$

Cette distance au noyau où la densité de probabilité est maximale est égale au rayon de l'orbite de l'électron dans le n^e état prévu par la théorie de Bohr (Tome I, éq. (6.40)). La formule de Stirling[62] montre que $R_{nn-1}(r_{\max}) \simeq \tilde{a}_0^{-3/2}\pi^{-1/4}n^{-11/4}$; le maximum $\rho_{nn-1}(r_{\max})$ décroît avec n comme $n^{-3/2}$ (voir la question 3 dans le problème 19.3.11).

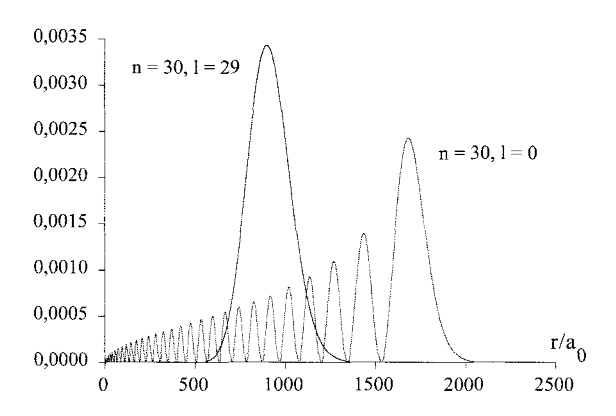


Figure 19.6: Densité radiale $\rho_{nl}(r) \stackrel{\text{déf}}{=} r^2[R_{nl}(r)]^2$ pour l'état lié $n = 30$, $l = 0$, et 29. En ordonnée est porté le produit $a_0^3\rho_{nl}$. Noter que pour $l = n-1 \equiv l_{\max}$, l'électron est relativement bien localisé autour du cercle de Bohr de rayon $n^2 a_0$.

Par ailleurs, comme $u(r)$ se comporte comme r^{l+1} près de $r = 0$, R_{nl} varie comme r^l : seuls les états s ($l = 0$) ont une densité de probabilité $[R(r)]^2$ non nulle à l'origine[63]. Pour $n \gg 1$, la densité radiale $r^2[R_{nl}(r)]^2$ part de $r = 0$ extrêmement lentement, le coefficient du terme en r^{2l+2} étant proportionnel à $n^{-2(l+2)}$; ceci traduit le fait que, dans les états fortement excités, la probabilité est quasi-nulle de trouver l'électron près du noyau. Les premières fonctions radiales normalisées selon (19.180) sont ($\rho = r/a_0$) :

$$R_{10}(r) = \frac{2}{\sqrt{a_0^3}}\,\mathrm{e}^{-\rho}\ , \quad R_{20}(r) = \frac{1}{\sqrt{2a_0^3}}\,(1-\frac{\rho}{2})\,\mathrm{e}^{-\rho/2}\ , \quad R_{21}(r) = \frac{1}{2\sqrt{6a_0^3}}\,\rho\,\mathrm{e}^{-\rho/2}\ , \tag{19.187}$$

$$R_{30}(r) = \frac{2}{3\sqrt{3a_0^3}}\,(1-\frac{2\rho}{3}+\frac{2\rho^2}{27})\,\mathrm{e}^{-\rho/3}\ , \quad R_{31}(r) = \frac{8}{27\sqrt{6}}\,\rho(1-\frac{\rho}{6})\,\mathrm{e}^{-\rho/3} \tag{19.188}$$

[62] $\mathrm{N} \gg 1$: $N! \simeq \sqrt{2\pi N}\,(N/\mathrm{e})^N$.

[63] Cette particularité est responsable de la capture K (aussi appelée *désintégration β inverse)*, permettant à un noyau d'absorber un électron situé dans la couche la plus interne ; le réarrangement conduit à la formation de l'isotope ${}_{Z-1}^{\ \ A}\mathrm{X}$ avec émission d'un neutrino ; un exemple classique :

$$^{26}_{13}\mathrm{Al} + \mathrm{e}^- \ \rightarrow\ ^{26}_{12}\mathrm{Mg} + \nu_e\ . \tag{19.186}$$

$$R_{32}(r) = \frac{4}{81\sqrt{30}}\,\rho^2\,\mathrm{e}^{-\rho/3} \ . \tag{19.189}$$

Pour un atome hydrogénoïde de charge nucléaire $Z|e|$) (en fait un ion $Z-1$ fois ionisé), il suffit de remplacer partout e'^2 par Ze'^2 ; l'énergie E_n devient Z^2E_n et la distribution électronique est contractée d'un facteur Z :

$$\boxed{|e| \rightarrow Z|e| \qquad \Longrightarrow \qquad E_n \rightarrow Z^2E_n \ , \quad \tilde{a}_0 \rightarrow \frac{\tilde{a}_0}{Z}} \tag{19.190}$$

Avec la forme $E_n = -\frac{\alpha^2}{2n^2}mc^2 \rightarrow -\frac{(Z\alpha)^2}{2n^2}mc^2$ et $\alpha^{-1} \simeq 137$, on voit que les effets relativistes deviennent vite pertinents pour $n \gtrsim 1$, même pour un élément assez léger.

Dans les fonctions d'onde complètes $\psi_{nlm}(r,\theta,\phi) = R_{nl}(r)\,Y_{lm}(\theta,\phi)$, les Y_{lm} portent toute la dépendance angulaire ; les états $l=0$ sont à symétrie *sphérique* (voir (17.98)) :

$$\mathcal{R}_{\vec{u},\,\Theta}\psi_{n00}(\vec{r}\,) \equiv \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\Theta\vec{u}.\vec{L}}\psi_{n00}(\vec{r}\,) = \psi_{n00}(\vec{r}\,) \ . \tag{19.191}$$

Au contraire, les états $l>0$ ont seulement une symétrie *axiale* en module, étant invariants en module par rotation autour de Oz. Plus précisément, et comme tous sont propres de L_z, toute rotation autour de Oz donne au plus un facteur de phase à chacun d'entre eux :

$$\mathcal{R}_{\mathrm{O}z,\,\Theta}\psi_{nlm}(\vec{r}\,) \equiv \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\Theta L_z}\psi_{nlm}(\vec{r}\,) = \mathrm{e}^{-\mathrm{i}m\Theta}\psi_{nlm}(\vec{r}\,) \ , \tag{19.192}$$

tandis qu'une rotation autour d'un axe quelconque mélange entre elles les $\{\psi_{nlm}(\vec{r}\,)\}_m$, n et l restant les mêmes. Il ne faut pas être choqué d'obtenir des états propres ayant une symétrie *inférieure* à celle du Hamiltonien : après tout, le choix de l'axe Oz, axe de quantification du moment cinétique, est totalement arbitraire et n'importe quel autre choix conviendrait tout autant. D'ailleurs, toute combinaison linéaire impliquant les $2l+1$ harmoniques sphériques $Y_{lm=0,\,\pm1,\,\pm2,\ldots,\pm l}$ est tout autant acceptable – puisqu'aucune direction n'est privilégiée, ce qui se traduit par la dégénérescence de l'énergie vis-à-vis de m, quelle que soit la forme précise du potentiel central $V(r)$; une telle combinaison arbitraire ne possède évidemment pas la symétrie axiale simple de l'*une* des Y_{lm}.

En ce qui concerne les états non-liés, la fonction radiale est donnée par l'expression (19.168). Si l'on choisit la normalisation conventionnelle :

$$\int_0^{+\infty} R_{El}(r)R_{E'l}(r)\,r^2\mathrm{d}r = \delta(E-E') \ , \tag{19.193}$$

$R_{El}(r)$ a pour dimension (énergie×longueur^3)$^{-1/2}$; la constante de normalisation C_{El} est[64] ([48], § 36) :

$$C_{El} = \sqrt{\frac{2\mu\kappa}{\pi\hbar^2}}\,\mathrm{e}^{\pi\nu/2}\frac{|\Gamma(l+1-\mathrm{i}\nu)|}{(2l+1)!} \ , \tag{19.194}$$

[64]Le calcul de C_{El}, assez technique, est détaillé dans la référence [48], Appendice, § f.

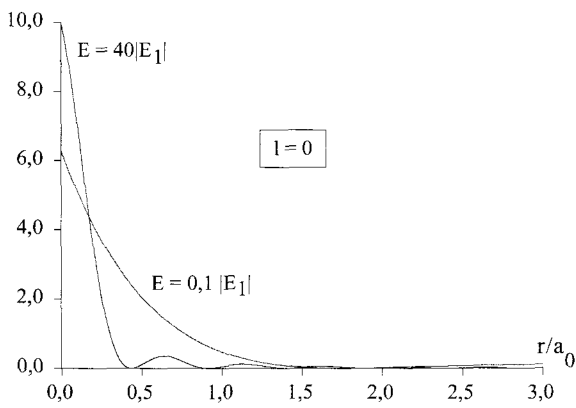


Figure 19.7: Fonction $|R_{El}(r)|^2$ pour $l = 0$ et deux états de basse et haute énergie. E_1 désigne l'énergie de l'état fondamental ; l'abscisse est le rapport r/a_0. En ordonnée est portée la quantité $\pi a_0^3 |E_{n=1}| \big(R_{El}(r)\big)^2$.

d'où :

$$\boxed{R_{El}(r) = \sqrt{\frac{2\mu\kappa}{\pi\hbar^2}}\, \mathrm{e}^{\pi\nu/2} \frac{|\Gamma(l+1-\mathrm{i}\nu)|}{(2l+1)!}\, (2\kappa r)^l\, \mathrm{e}^{-\mathrm{i}\kappa r}\Phi(l+1+\mathrm{i}\nu, 2l+2, 2\mathrm{i}\kappa r)} \qquad (19.195)$$

avec $E = \frac{\hbar^2\kappa^2}{2\mu} = \frac{e'^2}{2\nu^2\tilde{a}_0^2} = \frac{\mu e'^4}{2\nu^2\hbar^2}$. Il n'y a plus ici aucune contrainte entre ν et l – puisque l'inégalité $E > \min_r V(r)$ est automatiquement satisfaite quand $E > 0$ – de sorte que ces états sont infiniment dégénérés : pour chaque valeur de E (donc de ν), l'entier l peut prendre toutes les valeurs dans $\mathbb{N}$, et pour un l fixé, m varie par pas d'une unité entre $-l$ et $+l$. La fonction Φ pertinente pour les états non-liés a la représentation intégrale (voir problème 19.3.13, égalité (19.311)) :

$$\Phi(l+1+\mathrm{i}\nu,\, 2l+2,\, 2\mathrm{i}\kappa r) = \frac{(2l+1)!}{\Gamma(l+1+\mathrm{i}\nu)\Gamma(l+1-\mathrm{i}\nu)} \int_0^1 x^{l+\mathrm{i}\nu}(1-x)^{l-\mathrm{i}\nu}\, \mathrm{e}^{2\mathrm{i}\kappa r x}\, \mathrm{d}x \ . \qquad (19.196)$$

L'état non-lié de plus basse énergie, $E \to 0_+$, mérite l'attention ; dans cette limite, $\nu \to +\infty$ et $\kappa \to 0$. Sur l'expression en série entière de $\Phi(l+1+\mathrm{i}\nu, 2l+2, 2\mathrm{i}\kappa r)$, (19.142), on voit que chacun des termes de la série a une limite finie ; plus précisément, on obtient :

$$\lim_{E\to 0_+} \Phi(l+1+\mathrm{i}\nu, 2l+2, 2\mathrm{i}\kappa r) = \sum_{n\in\mathbb{N}} \frac{(-1)^n}{(2l+2)(2l+3)...(2l++n)\, n!}(2\rho)^n \ , \qquad (19.197)$$

avec $\rho = \mu e'^2 r/\hbar^2 = r/\tilde{a}_0$. On reconnaît les coefficients du développement[65] en série entière de la fonction de Bessel J_{2l+1}, au facteur $(2l+1)!$ près ; très précisément, on a :

$$\lim_{E\to 0_+} \Phi(l+1+\mathrm{i}\nu, 2l+2, 2\mathrm{i}\kappa r) = (2l+1)!(2\rho)^{-l-1/2}\, J_{2l+1}(\sqrt{8\rho}) \ ; \qquad (19.198)$$

[65] Voir éq. (19.76).

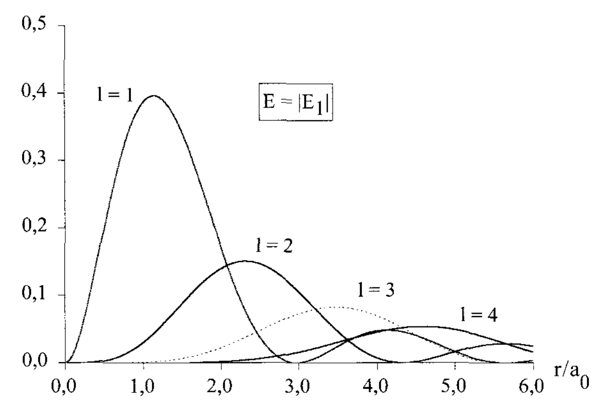


Figure 19.8: Fonction $|R_{El}(r)|^2$ pour $E = |E_1|$ et pour $l = 1,\ 2,\ 3,\ 4$. E_1 désigne l'énergie de l'état fondamental ; l'abscisse est le rapport r/a_0. En ordonnée est portée la quantité $\pi a_0^3 |E_{n=1}| \big(R_{El}(r)\big)^2$.

par ailleurs, la limite du produit $C_{\kappa l}(2\kappa r)^l$ s'obtient facilement en utilisant la relation fonctionnelle $\Gamma(z+1) = z\Gamma(z)$ et sachant que $|\Gamma(\mathrm{i}\nu)|^2 = \frac{\pi}{\nu \sinh \pi\nu}$. Finalement, on obtient :

$$\boxed{R_{E=0_+\,l}(r) = \sqrt{\frac{2m}{\hbar^2 r}}\, J_{2l+1}(\sqrt{8r/\tilde{a}_0})} \tag{19.199}$$

Pour $r/\tilde{a}_0 \gg 1,\ l$, la densité $|R_{E=0_+\,l}(r)|^2$ décroît donc comme r^{-2} en oscillant.

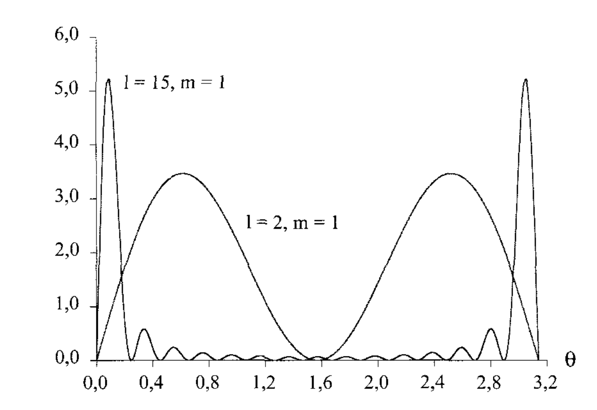


Figure 19.9: Variation en fonction de θ du courant azimutal $j_\phi(r,\ \theta)$ pour deux valeurs de l et pour des petites valeurs de $\langle L_z \rangle$ (unités arbitraires).

Pour terminer cette analyse, calculons le courant $\vec{j}$ dans un état stationnaire Ψ construit avec l'état propre lié ψ_{nlm}. On se souvient que Y_{lm} contient le facteur $\mathrm{e}^{im\phi}$

et qu'il s'agit de la seule contribution complexe à la fonction d'onde ; par ailleurs, le gradient a pour composantes sphériques :

$$\vec{\nabla} = \left(\frac{\partial}{\partial r},\, \frac{1}{r}\frac{\partial}{\partial \theta},\, \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}\right) \ . \tag{19.200}$$

Comme dans l'expression du courant on fait la différence entre un nombre complexe et son conjugué, seuls comptent les termes imaginaires purs – tous les autres se compensent deux à deux ; seule la dérivation en ϕ donne une contribution complexe. Au total, j_r et j_θ sont nuls et le vecteur $\vec{j}$ a pour seule composante non nulle j_ϕ, indiquant un courant annulaire tournant autour de Oz ; un calcul simple donne[66] :

$$j_\phi(r,\, \theta) = \frac{m\hbar}{\mu r\sin\theta}\,|\psi_{nlm}(r,\, \theta\, \phi)|^2 = \frac{(2l+1)(l-|m|)!}{4\pi(l+m)!}\,\frac{m\hbar}{\mu r\sin\theta}\,[R_{nl}(r)\,P_l^{|m|}(\cos\theta)]^2 \ . \tag{19.201}$$

Le fluide de probabilité tourne donc autour de Oz ; ce n'est pas une rotation rigide ("en bloc") – caractérisée par le fait que le courant j_ϕ est proportionnel à la distance à l'axe de rotation, soit $r\sin\theta$. Le moment cinétique global $\vec{K}$ lié à la rotation du fluide est visiblement dirigé le long de Oz, et se réduit à sa seule composante K_z :

$$\mathrm{d}\vec{K} = \mu\vec{r}\times\vec{j}\,\mathrm{d}^3r \tag{19.202}$$

et :

$$K_z = \mu\int_{\mathbb{R}^3} j_\phi\,\sin\theta\,\mathrm{d}^3r = \mu\int_{\mathbb{R}^3}\frac{m\hbar}{\mu}\,|\psi_{nlm}|^2\,\mathrm{d}^3r = m\hbar \ . \tag{19.203}$$

Autrement dit, le moment cinétique de rotation du courant de probabilité n'est autre que la composante L_z du moment cinétique orbital. Cet écoulement permanent est visiblement rendu possible parce que le système considéré est plongé dans $\mathbb{R}^3$. Il est nul dans un état sphérique (états s) et, à l quelconque, pour l'état ayant $m = 0$ ($\langle L_z\rangle = 0$). La variation angulaire de $j_\phi(r,\, \theta)$ en fonction de θ dépend beaucoup de la valeur de l, puisque $P_l^{|m|}$ s'annule $l-m$ fois dans $]-1,\, +1[$, et deux fois aux bornes (aux "pôles"). Si l est petit, $j_\phi(r,\, \theta)$ présente peu d'oscillations ; au contraire, si $l \gg 1$ et m petit ($\langle L_z\rangle$ est alors petit), $j_\phi(r,\, \theta)$ est une fonction très structurée de θ (voir fig. 19.9), présentant des *maxima* très pointus au voisinage immédiat des pôles Nord et Sud. À l'inverse, pour la valeur maximum de m, donc de $\langle L_z\rangle$, $j_\phi(r,\, \theta)$ a un maximum en $\theta = \pi/2$, très fin si $l \gg 1$ (voir fig 19.10), le courant annulaire ayant essentiellement la forme d'un disque équatorial, dont la densité décroît comme $|R_{nl}(r)|^2$ en fonction de la distance r au centre.

▼ À propos des polynômes de Laguerre

Il s'agit ici de donner quelques compléments techniques indispensables concernant la partie radiale $R(r)$ des états propres de l'atome hydrogène. Les polynômes de Laguerre

[66]La composante azimutale ne dépend pas de ϕ.

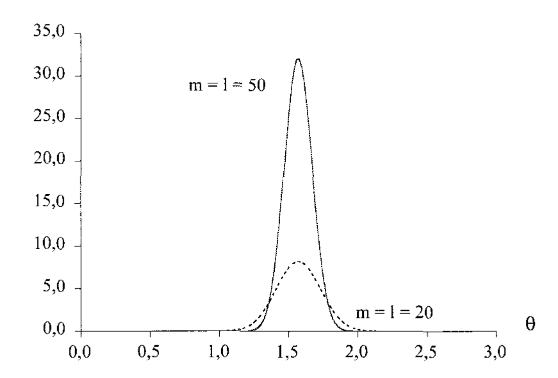


Figure 19.10: Variation en fonction de θ du courant azimutal $j_\phi(r,\,\theta)$ pour deux grandes valeurs de l et pour $m = l$, valeur maximale de $\langle L_z\rangle$ (unités arbitraires).

$L_p(z)$ ($p \in \mathbb{N}$) sont par définition[67], les solutions analytiques à l'origine de l'équation de Laplace particulière :

$$z\,\frac{\mathrm{d}^2 L_p}{\mathrm{d}z^2} + (1-z)\,\frac{\mathrm{d}L_p}{\mathrm{d}z} + p\,L_p = 0 \ ; \tag{19.204}$$

on achève de les définir en posant conventionnellement :

$$L_p(z=0) = p! \ , \tag{19.205}$$

de sorte que :

$$\boxed{L_p(z) = p!\,\Phi(-p,\,1,\,z)} \tag{19.206}$$

la fonction Φ étant définie en (19.142). $L_p(z)$ est un polynôme de degré p, dont on peut montrer que tous les zéros sont réels.

Par ailleurs, en dérivant q fois (19.204) membre à membre après y avoir remplacé p par $p+q$, il est facile de voir que la q^{e} dérivée de L_{p+q}, $L_{p+q}^{(q)}$, satisfait :

$$z\,\frac{\mathrm{d}^2 L_{p+q}^{(q)}}{\mathrm{d}z^2} + (q+1-z)\,\frac{\mathrm{d}L_{p+q}^{(q)}}{\mathrm{d}z} + p\,L_{p+q}^{(q)} = 0 \ ; \tag{19.207}$$

c'est l'équation différentielle (19.139) obtenue plus haut avec l'identification :

$$\lambda = 2n \ , \qquad q+1 = 2l+2 \ , \quad p = n-l-1 \ . \tag{19.208}$$

Ceci conduit aux polynômes *associés* de Laguerre, précisément définis comme suit :

$$\boxed{L_p^q(z) \stackrel{\text{déf}}{=} (-1)^q\,\frac{\mathrm{d}^q}{\mathrm{d}z^q}\,L_{p+q}(z)} \tag{19.209}$$

[67] Cette définition fixe les polynômes L_p à un facteur près, évidemment. Le choix conventionnel adopté ici, (19.205), est le plus usité, mais on trouve aussi la définition $L_p(0) = 1$.

qui permet d'ailleurs d'identifier $L_p^{q=0}(z)$ avec $L_p(z)$. $L_p^q(z)$, en tant que q^{e} dérivée d'un polynôme de degré $p+q$, est un polynôme de degré p, tout comme $L_p(z)$. La relation avec la fonction hypergéométrique est :

$$\boxed{L_p^q(z) = (-1)^q \frac{(p+q)!^2}{p!q!}\,\Phi(-p,\,q+1,z)} \tag{19.210}$$

Compte tenu de l'identification (19.208), la partie radiale $R_{nl}(r)$ implique donc le polynôme associé $L_{n-l-1}^{2l+1}(2k_n r)$, k_n étant défini en (19.154).

Les relations étant précisées entre R_{nl} et les polynômes L_p^q, et entre les L_p^q et les L_p, il est utile d'en dire un peu plus sur ces derniers. La résolution de l'équation (19.204) peut être entreprise par la méthode de Laplace, qui fonctionne efficacement pour toute équation différentielle dont les coefficients sont au plus des fonctions linéaires de la variable. Cette méthode consiste à poser la fonction inconnue sous la forme :

$$L_p(z) = \int_C \mathrm{d}s\, \mathrm{e}^{-sz} f(s) \ , \tag{19.211}$$

où C est un contour (éventuellement fermé) dans $\mathbb{C}$ à préciser ultérieurement, f devenant la fonction inconnue. Le résultat d'une telle transformation est de *réduire* d'une unité l'ordre de l'équation différentielle ; comme celle-ci est initialement d'ordre deux, on est ramené à une quadrature[68]. La terminologie est évidente : l'intégrale (19.211) ressemble manifestement à une transformation de Laplace ordinaire (pour plus de détails, voir [48], Appendice, § a, et problème 19.3.13 ci-après). C'est ainsi que l'on peut établir l'expression suivante :

$$\boxed{L_p(z) = \frac{p!}{2\mathrm{i}\pi} \oint \mathrm{d}s\, \frac{\mathrm{e}^{-sz}}{s^{p+1}}\,(1+s)^p} \tag{19.212}$$

où l'intégrale est prise le long d'une boucle entourant une fois et une seule l'origine du plan de s et parcourue dans le sens positif. Avec (19.209), on en déduit l'expression intégrale des polynômes associés :

$$\boxed{L_p^q(z) = \frac{(p+q)!}{2\mathrm{i}\pi} \oint \mathrm{d}s\, \frac{\mathrm{e}^{-sz}}{s^{p+1}}\,(1+s)^{p+q}} \tag{19.213}$$

Avec la représentation (19.212), on peut écrire :

$$L_p(z) = (-1)^p\,\frac{p!}{2\mathrm{i}\pi}\,\mathrm{e}^z\,\frac{\mathrm{d}^p}{\mathrm{d}z^p} \oint \mathrm{d}s\, \frac{\mathrm{e}^{-(s+1)z}}{s^{p+1}} = (-1)^p\,\frac{p!}{2\mathrm{i}\pi}\,\mathrm{e}^z\,\frac{\mathrm{d}^p}{\mathrm{d}z^p} \oint \mathrm{d}s\, \frac{\mathrm{e}^{-sz}}{s^{p+1}}\,\mathrm{e}^{-z} \ ; \tag{19.214}$$

en développant maintenant l'exponentielle e^{-sz} sous l'intégrale, et en appliquant le théorème des résidus (la convergence est visiblement uniforme) – ou en appliquant la formule de Cauchy –, il vient simplement :

$$\boxed{L_p(z) = \mathrm{e}^z\,\frac{\mathrm{d}^p}{\mathrm{d}z^p}\left(z^p\mathrm{e}^{-z}\right)} \tag{19.215}$$

[68] Un autre cas où la méthode de Laplace est très utile est celui de l'équation d'Airy, que l'on rencontre en tant qu'équation aux valeurs propres pour une particule dans un champ constant (voir Tome I, p. 339).

qui est, pour les polynômes de Laguerre, une formule *à la* Rodrigues analogue à celle obtenue pour ceux de Legendre (voir chapitre 18, éq. (18.103)). L'exponentielle en facteur assure bien que L_p est un simple polynôme, de degré p, dont le coefficient du terme de plus haut degré est d'ailleurs égal à $(-1)^p$. Le calcul de la dérivée p^e par la formule de Leibniz[69, 70] donne sans difficulté l'expression suivante :

$$\boxed{L_p(z) = \sum_{r=0}^{p} \frac{(-1)^r\, p!^2}{(p-r)!\, r!^2}\, z^r} \tag{19.217}$$

Comme d'habitude, il est utile de définir la fonction génératrice des polynômes d'intérêt ; celle des L_p s'écrit par définition :

$$F(t, z) \stackrel{\text{déf}}{=} \sum_{p=0}^{+\infty} L_p(z)\, \frac{t^p}{p!} \; ; \tag{19.218}$$

le rayon de convergence de cette série entière se déduira ([49], § 7.1) de la forme resommée (voir (19.225)). $F(t, z)$ est donc la fonction dont les coefficients de la série de Taylor autour de $t = 0$ sont[71] les $L_p(z)$:

$$L_p(z) = \left(\frac{\partial^p}{\partial t^p}\, F(t, z)\right)_{t=0} \; . \tag{19.220}$$

$F(t, z)$ s'obtient en partant de l'expression (19.212), où l'on développe le binôme $(1+s)^p$. En utilisant la convergence uniforme des doubles séries et le théorème des résidus, on trouve :

$$F(t, z) = \sum_{r=0}^{+\infty} \frac{(-zt)^r}{r!} \sum_{p=0}^{+\infty} \mathrm{C}_{p+r}^{p} t^p \; . \tag{19.221}$$

La sommation interne s'écrit :

$$\sum_{p=0}^{+\infty} \mathrm{C}_{p+r}^{p} t^p = \sum_{p=0}^{+\infty} \frac{t^p}{p!} \frac{(p+r)!}{r!} = \sum_{p=0}^{+\infty} (p+r)(p+r-1)\ldots(r+2)(r+1)\, \frac{t^p}{p!} \; , \tag{19.222}$$

[69]Gottfried Wilhelm von Leibniz (1646-1716) philosophe, physicien et mathématicien, a apporté des contributions décisives en analyse, notamment en calcul infinitésimal (on lui attribue généralement la paternité de la notion de dérivée et d'intégrale). Vers 1675, une vive controverse éclata entre lui et Newton au sujet de l'invention du calcul infinitésimal.

[70]Pour deux fonctions f et g, la dérivée n^e du produit est :

$$(fg)^{(n)} = \sum_{p=0}^{n} \mathrm{C}_{\mathrm{n}}^{\mathrm{p}} f^{(n-p)} g^{(p)} \; . \tag{19.216}$$

[71]De ce fait, par la formule de Cauchy, on a :

$$L_p(z) = \frac{p!}{2\mathrm{i}\pi} \oint \frac{F(\xi, z)}{\xi^{p+1}}\, \mathrm{d}\xi \; , \tag{19.219}$$

où le contour est une petite boucle autour de l'origine. En injectant l'expression (19.225) de $F(t, z)$ et en posant $\frac{\xi}{1-\xi} = s$, on retombe bien sur l'expression intégrale (19.212) de $L_p(z)$.

qui ressemble à une série de Taylor ; de fait, en observant que :

$$\frac{\mathrm{d}^p}{\mathrm{d}t^p}\frac{1}{(1-t)^{r+1}} = \frac{(r+1)(r+2)\ldots(r+p)}{(1-t)^{r+1+p}} \quad , \tag{19.223}$$

on voit que :

$$(r+1)(r+2)\ldots(r+p) = \left(\frac{\mathrm{d}^p}{\mathrm{d}t^p}\frac{1}{(1-t)^{r+1}}\right)_{t=0} \; ; \tag{19.224}$$

ainsi, la somme interne n'est autre que la série de Taylor de $(1-t)^{-(r+1)}$ autour de l'origine. D'où, reportant dans (19.221) :

$$\boxed{F(t,z) = \sum_{r=0}^{+\infty}\frac{(-zt)^r}{r!}\frac{1}{(1-t)^{r+1}} = \frac{1}{1-t}\,\mathrm{e}^{-zt/(1-t)}} \; . \tag{19.225}$$

La série entière (19.218) ne converge que si $|t| < 1$, mais sa somme peut être prolongée analytiquement et devient ainsi une fonction analytique dans le plan complexe ouvert privé du point $t = 1$, où elle a une singularité essentielle tant que $z \neq 0$. Sur cette dernière expression de $F(t,z)$, et par comparaison avec la définition (19.218), on vérifie bien la valeur conventionnelle $L_p(0) = p!$:

$$F(t, z=0) = \frac{1}{1-t} = \sum_{p=0}^{+\infty} t^p = \sum_{r=0}^{+\infty} L_p(0)\,\frac{t^p}{p!} \qquad (|t| < 1) \; . \tag{19.226}$$

Le polynôme associé, L_p^q, est la dérivée q^{e} d'un polynôme de degré $p+q$, c'est donc encore un polynôme de degré p. En raison de la définition (19.209) des L_p^q, leur fonction génératrice est $F_q(t,z)$:

$$F_q(t,z) \stackrel{\text{déf}}{=} \sum_{p=0}^{+\infty} L_p^q(z)\frac{t^p}{(p+q)!} = \sum_{p=0}^{+\infty}\frac{\mathrm{d}^q}{\mathrm{d}z^q}L_{p+q}(z)\frac{(-1)^q t^p}{(p+q)!} = (-t)^{-q}\frac{\mathrm{d}^q}{\mathrm{d}z^q}\sum_{p=q}^{+\infty}L_p(z)\frac{t^p}{p!} \; . \tag{19.227}$$

La somme n'est pas exactement $F(t,z)$, mais comme les termes manquants sont des monômes en z de degré variant de 0 à $q-1$, leur q^{e} dérivée est nulle, et on peut de fait commencer la sommation à $p = 0$. Au total :

$$F_q(t,z) = (-t)^{-q}\,\frac{\mathrm{d}^q}{\mathrm{d}z^q}\sum_{p=0}^{+\infty} L_p(z)\,\frac{t^p}{p!} = (-t)^{-q}\,\frac{\partial^q}{\partial z^q}\,F(t,z) \; . \tag{19.228}$$

soit :

$$\boxed{F_q(t,z) = \frac{1}{(1-t)^{q+1}}\,\mathrm{e}^{-zt/(1-t)}} \tag{19.229}$$

Faisant $z = 0$, on en déduit :

$$F_q(t, z=0) = \sum_{p=0}^{+\infty} L_p^q(0)\,\frac{t^p}{(p+q)!} = \frac{1}{(1-t)^{q+1}} \; ; \tag{19.230}$$

reprenant alors le dévelopement de Taylor de $(1-t)^{-(q+1)}$ (voir (19.224)), on voit que :

$$\boxed{L_p^q(z=0) = \frac{[(p+q)!]^2}{p!q!}} \tag{19.231}$$

Bien sûr, à l donné, les polynômes associés doivent satisfaire des relations d'orthogonalité puisqu'ils entrent dans l'expression des fonctions propres du Hamiltonien radial H_{eff}. De fait, on doit avoir :

$$\int_0^{+\infty} r^{2l+2} L_{n-l-1}^{2l+1}(2k_n r) L_{n'-l-1}^{2l+1}(2k_{n'} r)\, \mathrm{e}^{-\,(k_n+k_{n'})r}\, \mathrm{d}r = 0 \quad \text{si } n \neq n' \ . \tag{19.232}$$

En revanche, deux fonctions R_{nl} et $R_{nl'}$ n'ont nullement à être orthogonales[72] (et ne le sont pas). Le problème 19.3.12 reprend et complète ces éléments techniques. ▲

19.2.3 Symétrie dynamique du potentiel Coulombien

La section 17.6 a présenté brièvement le lien indissoluble entre symétrie et dégénérescence, imposé par une évidence physique : l'énergie d'un système ne saurait varier quand on effectue une opération de symétrie (géométrique ou abstaite) qui, par nature, laisse le système totalement indifférent. Dans le cas précis de l'atome d'hydrogène, la symétrie sphérique dans $\mathbb{R}^3$ permet de prévoir avec certitude que l'énergie ne dépend pas du nombre quantique magnétique m, puisque l'axe de quantification Oz est complètement arbitraire, et que l'atome se moque de savoir quel axe est conventionellement choisi pour conduire les calculs. À nouveau, cette symétrie propre est induite par la symétrie galiléenne affirmant l'isotropie de l'espace.

Le calcul de l'expression de l'énergie de l'atome, produisant l'expression (19.156), provoque inévitablement la surprise, puisque le nombre quantique l n'y figure pas : s'il était d'avance acquis que l'énergie ne peut dépendre de m, il apparaît qu'elle ne dépend pas non plus de la valeur du module du moment cinétique. Aucun argument de symétrie *évidente* ne permet de prévoir cette propriété, et ce fait est sans doute à l'origine de la terminologie (historique) ayant consacré l'expression *dégénérescence accidentelle.* Comme on va le voir, il n'y a en fait aucun *accident* dans tout cela, grâce à l'existence d'une symétrie fort subtile dont la mise en évidence a été quelque peu tardive.

Avant d'expliquer les fondements de cette symétrie, il est bon d'ajouter que l'atome d'hydrogène – c'est-à-dire en fait le champ coulombien nu – n'est pas un cas unique [46], même si les systèmes ayant une telle singularité sont relativement exception-

[72] En revanche, les deux fonctions propres *complètes* ψ_{nlm} et $\psi_{nl'm}$ sont bien sûr orthogonales, grâce à $\langle Y_{lm} | Y_{l'm} \rangle = \delta_{ll'}$.

nels. Un autre exemple est l'oscillateur harmonique isotrope[73,74] dans $\mathbb{R}^D$, où $D \geq 2$; en effet, il s'agit à nouveau d'un système ayant la symétrie (propre) sphérique, et l'on peut à nouveau prévoir que les niveaux d'énergie ne dépendront pas de la valeur de la projection du moment cinétique attachée à m. En réalité, la dégénérescence révélée par le calcul est bien plus grande, ce dont on peut être convaincu en quelques mots.

Prenons le cas d'un oscillateur bidimensionnel ($D = 2$), le plan étant fixé par des contraintes physiques bien définies (la valeur initiale du moment cinétique, par exemple) ; un tel système possède dès lors la seule symétrie de rotation autour de l'axe perpendiculaire au plan (noté Oz). Traité en coordonnées polaires du plan (r, θ), un ECOC est formé par (H, $L_z = -\mathrm{i}\hbar\partial_\theta$) et la symétrie conduit inévitablement à une dégénérescence d'ordre *deux*, correspondant aux deux sens de rotation possibles n'affectant pas l'énergie ; les états propres seront des fonctions $\psi_{n_r \pm\,\mathrm{pol}}(r,\,\theta)$ de la forme $f_{n_r}(r)\mathrm{e}^{\pm\mathrm{i}\theta}$, l'énergie[75] étant une fonction du seul nombre quantique n_r, les $f_{n_r}(r)$ apparaissant naturellement lors de la résolution de l'équation radiale associée au Hamiltonien bidimensionnel :

$$H_{2\,\mathrm{pol}}(r\,\theta) \stackrel{\text{déf}}{=} \frac{\vec{p}^{\,2}}{2\mu} + \frac{1}{2}m\omega^2 r^2 \quad , \tag{19.233}$$

avec l'opérateur $\vec{p}^{\,2}$ exprimé en polaires dans $\mathbb{R}^2$. Nantie des bonnes conditions aux limites, la résolution complète de cette équation révèle la quantification spontanée des états propres acceptables physiquement suivant le scénario habituel.

En fait, le même problème peut tout autant être résolu en coordonnées cartésiennes (x, y) du plan, en écrivant cette fois le Hamiltonien :

$$H_{2\,\mathrm{cart}}(x,\,y) \stackrel{\text{déf}}{=} \frac{\vec{p}^{\,2}}{2\mu} + \frac{1}{2}m\omega^2(x^2 + y^2) \quad ; \tag{19.234}$$

avec $\vec{p}^{\,2} = p_x^2 + p_y^2$; la séparabilité des variables saute à nouveau aux yeux, puisque $H_{2\,\mathrm{cart}}(x,\,y)$ est de la forme $H_1(x)+H_1(y)$, $H_1(u)$ étant le Hamiltonien d'un oscillateur-1d

[73] À ce propos, il convient de faire le rapprochement avec un théorème dû à Bertrand, concernant l'existence d'orbites *fermées* pour un système dynamique classique. Bertrand (1873) a montré que, pour un point matériel ayant l'énergie potentielle $V(r)$ dérivable jusqu'au troisième ordre inclus, la condition nécessaire et suffisante pour que les orbites soient fermées est que le potentiel soit en r^2 ou en $-r^{-1}$ ([50], p. 247, [51]) – si on abandonne l'hypothèse du champ *central*, il existe d'autres formes de $V(\vec{r})$, plus ou moins exotiques, donnant également des orbites fermées. Il est intéressant de savoir que le cas harmonique bidimensionnel et le problème de Kepler se correspondent l'un dans l'autre par la transformation conforme $z \longrightarrow z^2$ accompagnée d'un *scaling* du temps (Goursat).

On se souvient que le traitement relativiste de Sommerfeld (Tome I, section 7.3.3) lève une certaine dégénérescence et produit la structure fine de la raie H_α ; or, dans ce cas précisément, les orbites ne sont plus fermées mais sont des *rosettes*, que l'on peut se représenter approximativement comme des ellipses tournant très lentement autour de l'un des foyers (voir Tome I, fig. 7.4 et fig. 32 dans le corrigé du problème 7.4.13).

[74] Joseph Louis François Bertrand (1822-1900) est connu pour ses travaux en analyse (par exemple, les séries de Bertrand) et en théorie des nombres. Il formula en 1845 une fameuse conjecture (*Postulat de Bertrand*) – démontrée cinq ans plus tard par Tchebychev –, affirmant qu'il existe au moins un nombre entier entre N et $2N-2$ quel que soit $N > 3$. Bertrand est aussi connu pour le *Paradoxe de Bertrand* en théorie des probabilités, tendant à montrer qu'une même question peut recevoir plusieurs réponses différentes (en réalité, il n'y a pas de paradoxe : les différentes formulations ne sont pas équivalentes, mais correspondent à des protocoles expérimentaux différents, d'où des réponses différentes). Bertrand fut le traducteur en français d'une partie des travaux de Gauss.

[75] Il n'y a ici que des états liés puisque $V(r \to \infty) \to +\infty$.

le long de l'axe Ou. En conséquence, ses fonctions propres $\psi_{n_x n_y\,\mathrm{cart}}(x,\,y)$ sont le simple produit $\psi_{n_x}(x)\psi_{n_y}(y)$ où $\psi_{n_u}(u)$ est la fonction propre (connue) de l'oscillateur linéaire le long de Ou ; la fonction $\psi_{n_x n_y\,\mathrm{cart}}(x,\,y)$ est associée à l'énergie propre $E_{n_x n_y} = E_{n_x} + E_{n_y}$ avec $E_{n_u} = (n_u + \frac{1}{2})\hbar\omega$. La dégénérescence est évidente : il y a un certain nombre de façons distinctes de choisir n_x et n_y de sorte que leur somme soit égale à un entier $N \geq 0$ donné, nombre visiblement égal à $N + 1$. Ainsi, dans le traitement en coordonnées cartésiennes, il saute aux yeux que la dégénérescence de l'oscillateur à $D = 2$ est *supérieure* à celle que prévoit un argument basé sur la seule symétrie sphérique quand on décide de décrire le même système en coordonnées polaires. La conclusion première est de dire que, dans cette dernière description, il existe une dégénérescence *accidentelle.* En fait, là encore, il existe une autre symétrie *cachée* à première vue, et qu'elle soit tout autant présente pour un oscillateur dans $\mathbb{R}^{D>2}$ ne fait pas l'ombre d'un doute : l'argument ci-dessus a été donné dans le cas simple $D = 2$ mais, visiblement, n'est nullement tributaire de ce choix particulier. On connaît encore d'autres cas de dégénérescence additionnelle : la particule dans une boîte plongée dans $\mathbb{R}^D$, le rotateur rigide, ... En suivant le même argument que celui déployé à propos de la particule libre (section 19.1.3, notamment ce qui justifie d'écrire le développement (19.88), p. 793), on voit que l'un quelconque des N_a états propres dégénérés ψ_{n_a} obtenus dans une représentation admet un développement sur les N_α états propres dégénérés ψ_{ν_α} de l'autre représentation, ces états ayant tous une seule et même énergie.

On en vient d'ailleurs à soupçonner que la dégénérescence additionnelle est étoitement liée au fait qu'il existe plusieurs jeux de coordonnées dans lesquelles le problème est séparable : pour l'oscillateur à $D = 2$, il s'agit des coordonnées polaires vs. cartésiennes ; pour l'atome d'hydrogène, la séparabilité se produit en coordonnées polaires *et* en coordonnées paraboliques. Cette suspicion est légitime, comme l'a montré Bargmann [52].

Ayant en tête le lien indissoluble entre symétrie et dégénérescence, l'apparition d'une dégénérescence *accidentelle* fait pressentir l'existence d'un groupe de symétrie plus vaste – au sens de l'inclusion des ensembles – que celui imposé de prime abord par l'évidence[76]. Pauli [44] semble être le premier à avoir réellement exploré cette voie, poursuivie peu après par Fock[77] [45]. C'est cette idée qui est maintenant développée pour le cas du champ Coulombien.

Classiquement, on sait que le potentiel Coulombien donne lieu à une constante du

[76]Cette symétrie est aussi parfois appelée *symétrie dynamique* car elle résulte d'une forme très particulière de l'énergie potentielle, et disparaît au moindre écart par rapport à cette loi : pour le potentiel Coulombien pur, en $1/r$, tout terme additif (effet d'écran – même à longue portée –, répulsion à courte distance, etc.) supprime la dégénérescence accidentelle (voir problème 19.3.15).

[77]Vladimir Aleksandrovich Fock (1898-1974) est l'auteur de contributions magistrales en Théorie quantique. On lui doit notamment la notion d'*espace de Fock*, essentielle pour le traitement des assemblées de particules identiques dont le nombre n'est pas fixé, et espace fondamental de la Théorie quantique des champs ; on peut se représenter cet espace comme un mille-feuilles, chaque feuillet étant un secteur correspondant à un nombre donné de particules, la variation de ce nombre traduisant des transitions entre les différents "plans" (voir chapitre 22, section 22.5). En collaboration avec Douglas Rayner Hartree (1897-1958), mathématicien et physicien, il inventa la méthode dite *de Hartree - Fock* pour le traitement des particules en interaction (voir chapitre 22, sous-section 23.1.3). Cette méthode est d'usage universel en théorie du Problème à N-corps, à toute échelle d'énergie, pour la détermination (approchée) de l'état fondamental d'un système.

mouvement remarquable, le vecteur dit de Lenz[78] - Runge[79] $\vec{R}$ dont la définition est :

$$\boxed{\vec{R}_{\rm cl} \stackrel{\text{déf}}{=} \frac{1}{\mu}\vec{p}\times\vec{L} - e'^2\,\frac{\vec{r}}{r}} \tag{19.235}$$

En effet, il n'est pas difficile de montrer que les crochets de Poisson de n'importe laquelle des composantes de ce vecteur avec $H \stackrel{\text{déf}}{=} \frac{p^2}{2m} - \frac{e'^2}{r}$ sont nuls. Un calcul simple permet en outre d'établir l'égalité (voir exercice 19.3.14) :

$$\vec{R}_{\rm cl}^2 = e'^4 + \frac{2}{\mu}H\vec{L}^2 \ . \tag{19.236}$$

Classiquement, la constance de ce vecteur a une interprétation géométrique simple : dirigé le long de l'axe principal de l'ellipse (pour un mouvement lié), sa longueur est égale à $[e'^4+(2/\mu)E\vec{L}^2]^{1/2}$, E désignant l'énergie (négative) du mouvement lié. Par comparaison avec l'expression (7.122) du Tome I, on voit que $\|\vec{R}\| = \varepsilon e'^2$, où ε est l'excentricité de l'ellipse. En définitive, dire que ce vecteur est constant c'est dire que l'axe focal ne varie pas au cours du temps, pas plus que l'excentricité. De toute évidence, on a aussi :

$$\vec{R}_{\rm cl}.\vec{L} = 0 \ , \tag{19.237}$$

puisque le premier membre est la somme de produits mixtes du genre $(\vec{A}\times\vec{B}).\vec{C}$ avec à chaque fois $\vec{C} = \vec{A}$ ou $\vec{B}$. L'orthogonalité de $\vec{R}_{\rm cl}$ et $\vec{L}$ est une évidence physique, puisque la trajectoire est dans le plan orthogonal à $\vec{L}$.

L'opérateur quantique $\vec{R}$ se forme en symétrisant comme d'habitude à partir des grandeurs fondamentales $\vec{r}$ et $\vec{p}$: $\vec{p}\times(\vec{r}\times\vec{p}) \rightarrow \frac{1}{2}\big(\vec{p}\times(\vec{r}\times\vec{p}) + (\vec{p}\times\vec{r})\times\vec{p}\big)$, soit :

$$\boxed{\vec{R} \stackrel{\text{déf}}{=} \frac{1}{2\mu}(\vec{p}\times\vec{L} - \vec{L}\times\vec{p}) - e'^2\,\frac{\vec{r}}{r}} \tag{19.238}$$

puisque l'égalité $\vec{r}\times\vec{p} = -\vec{p}\times\vec{r}$ reste vraie pour les deux opérateurs $\vec{r}$ et $\vec{p}$, quelle que soit la représentation (-q ou -p) choisie (voir la note 1 p. 718). Les opérateurs $\vec{R}$ et $\vec{L}$ ne commutent pas entre eux ; les commutateurs correspondants s'expriment commodément à l'aide du tenseur complètement antisymétrique ε_{uvw} qui, par définition, permet d'écrire commodément les relations caractéristiques de commutation d'un moment cinétique : $[L_u,\, L_v] = \mathrm{i}\hbar\,\varepsilon_{uvw}\,L_w$ (voir (18.4)) ; on trouve :

$$[L_u,\, R_v] = \mathrm{i}\hbar\,\varepsilon_{uvw}\,R_w \ , \qquad [R_u,\, R_v] = -2\,\frac{\mathrm{i}\hbar}{\mu}\,\varepsilon_{uvw}\,L_w\,H \ . \tag{19.239}$$

$$[L_u^2,\, R_v] = \mathrm{i}\hbar\,\varepsilon_{uvw}\,(L_uR_w + R_wL_u) \ , \qquad \vec{L}.\vec{R} = \vec{R}.\vec{L} = 0 \ . \tag{19.240}$$

[78]Heinrich Friedrich Emil Lenz (1804-1865) est surtout connu pour la formulation de la célèbre loi de l'Électromagnétisme portant son nom (1833). Il découvrit indépendamment (1842) la loi de Joule.

[79]Carl David Tolmé Runge (1856-1927) était mathématicien et physicien, et a aussi beaucoup contribué au développement de la spectroscopie. En collaboration avec Kutta (1867-1944), il est l'auteur de la méthode dite de *Runge - Kutta* pour la résolution approchée d'équations différentielles, universellement employée dans les traitements numériques. Un cratère de la Lune porte son nom.

En outre, bien sûr :

$$[H,\, \vec{R}\,] = 0 \tag{19.241}$$

puisque $\vec{R}$ est une constante du mouvement. La première égalité (19.239) est une illustration du fait que si K_1 et K_2 sont des constantes du mouvement, leur commutateur $[K_1,\, K_2]$ en est aussi une, en raison de l'identité de Jacobi (voir Tome I, éq. (13.188)). Par ailleurs, puisque $[R_u,\, L_u] = 0$, l'égalité (19.237) se transcrit à l'identique en termes d'opérateurs, $\vec{R}.\vec{L} = \vec{L}.\vec{R} = 0$; enfin, un calcul un peu laborieux permet d'établir l'égalité[80] :

$$\boxed{\vec{R}^{\,2} = e'^4 + \frac{2}{\mu}H(\vec{L}^{\,2} + \hbar^2)} \tag{19.242}$$

Avant de continuer, un commentaire s'impose au vu de cette expression. En effet, si elle ne coïncide pas avec l'expression classique (19.236) hâtivement symétrisée en tirant profit du fait que H et $\vec{L}^{\,2}$ commutent, il ne s'agit nullement d'une mise en défaut de la prescription de Schrödinger pour former les opérateurs quantiques à partir de leurs expressions classiques. La règle de Schrödinger concerne les deux grandeurs fondamentales $\vec{r}$ et $\vec{p}$; c'est sur les expressions classiques *explicitées* à l'aide de ces dernières qu'il convient d'effectuer les symétrisations adéquates[81]. Cette procédure conduit bien à la forme (19.242) qui, comparée à sa version classique (19.236), contient le terme additionnel $\hbar^2$... lequel est bien nul à la limite classique.

Cela étant, on peut voir immédiatement que l'existence de cette constante du mouvement propre au champ Coulombien conduit à une dégénérescence de l'énergie vis-à-vis de l. En effet, par (19.241), on a :

$$\langle n\, l\, m\,|\, [H,\, R_z]\, |n\, l'\, m\,\rangle = 0 \ . \tag{19.244}$$

Appelons provisoirement E_{nl} l'énergie d'un état propre de H ; (19.244) donne :

$$(E_{nl} - E_{nl'})\, \langle n\, l\, m\,|\, R_z\, |n\, l'\, m\,\rangle = 0 \ . \tag{19.245}$$

Comme les deux opérateurs R_z et $\vec{L}^{\,2}$ ne commutent pas, on peut affirmer que sur la base propre de $(H, \vec{L}^{\,2},\, L_z)$ – où notamment $\vec{L}^{\,2}$ est représenté par une matrice diagonale –, la matrice de R_z n'est pas diagonale ; ses éléments non-diagonaux sont donc, en général, non nuls. Il résulte alors nécessairement de (19.245) que :

$$\boxed{E_{nl} = E_{nl'}} \tag{19.246}$$

ce qui établit l'indépendance de l'énergie vis-à-vis de l. En définitive, la dégénérescence particulière vient du fait qu'il existe une constante du mouvement, $\vec{R}$, qui ne commute pas[82] avec $\vec{L}^{\,2}$, L_z. Deux ECOC sont ainsi disponibles : $(H,\, \vec{L}^{\,2},\, L_z)$ ou $(H,\, \vec{R}^{\,2},\, R_z)$,

[80] Dans ces écritures, on omet l'opérateur identité **1** en facteur des scalaires, afin d'alléger la notation.

[81] Par exemple, avec des opérateurs on a :

$$\vec{p} \times \vec{L} \equiv \vec{p} \times (\vec{r} \times \vec{p}) = \vec{p}^{\,2}\, \vec{r} - \vec{p}\,(\vec{p}.\vec{r}) \tag{19.243}$$

de sorte que l'opérateur hermitique $\vec{p} \times \vec{L} - \vec{L} \times \vec{p}$ est égal à $\vec{p}^{\,2}\, \vec{r} - \vec{p}\,(\vec{p}.\vec{r}) + \vec{r}\,\vec{p}^{\,2} - (\vec{r}.\vec{p})\,\vec{p}$.

[82] Tout comme la dégénérescence de rotation est liée à l'existence de deux autres constantes du mouvement, L_x et L_y, qui ne commutent pas avec L_z, lequel est membre de l'ECOC $(H,\, \vec{L}^{\,2},\, L_z)$.

correspondant aux deux jeux de coordonnées (sphériques ou paraboliques) pour lesquelles les coordonnées (variables) se séparent, i.e. permettent de factoriser les états propres.

Les trois composantes de $\vec{L}$ sont les générateurs infinitésimaux des rotations dans $\mathbb{R}^3$ et forment une algèbre[83] fermée ; l'existence d'une plus grande dégénérescence conduit tout naturellement à se poser la question de l'identification d'une algèbre "plus grande", incluant les composantes de $\vec{R}$, associée à un groupe de transformations à identifier. Dans ce but, et en se restreignant désormais à un sous-espace engendré par des états liés ayant tous la même énergie $E < 0$, il se révèle utile de définir un autre vecteur $\vec{R}\,'$ par :

$$\boxed{\vec{R}\,' \stackrel{\text{déf}}{=} \sqrt{\frac{\mu}{-2E}}\,\vec{R}} \tag{19.247}$$

Commençons par une analyse un peu formelle (voir [53], p. 234). Les différentes grandeurs dynamiques peuvent être renumérotées comme suit :

$$\vec{r} = (r_1,\, r_2,\, r_3)\ , \qquad \vec{p} = (p_1,\, p_2,\, p_3)\ , \tag{19.248}$$

$$\vec{L} = (L_x,\, L_y,\, L_z) \equiv (L_{23},\, L_{31},\, L_{12})\ . \tag{19.249}$$

Dans ces notations, on a les relations compactes :

$$L_{ij} = r_i p_j - r_j p_i\ , \qquad [r_i,\, p_j] = \mathrm{i}\hbar\,\delta_{ij}\ , \tag{19.250}$$

avec $i, j = 1, 2, 3$. Prolongeons maintenant ces deux dernières équations en "inventant" une quatrième dimension (coordonnée) r_4 et le moment correspondant p_4, et posons :

$$R'_x = L_{14}\ , \qquad R'_y = L_{24}\ , \qquad R'_z = L_{34}\ . \tag{19.251}$$

Il est alors facile de vérifier que les 6 quantités L_{ij} satisfont :

$$[L_{ij},\, L_{kl}] = \mathrm{i}\hbar\,(L_{li}\delta_{jk} + L_{kj}\delta_{il} + L_{ik}\delta_{jl} + L_{jl}\delta_{ik})\ , \tag{19.252}$$

montrant que les L_{ij} forment une algèbre fermée. Cette équation constitue aussi une écriture compacte, pour ces 6 générateurs, des relations de commutation des composantes du moment cinétique et du vecteur de Lenz-Runge. On peut alors concevoir que, tout comme L_x, L_y et L_z sont des générateurs des rotations dans $\mathbb{R}^3$, groupe désigné par SO(3), les 6 générateurs ainsi définis constituent ceux d'un groupe d'opérations dans un espace plus vaste ; il s'agit en fait du groupe des rotations dans $\mathbb{R}^4$, noté[84] SO(4). Bien sûr, la symétrie correspondante n'est pas de nature géométrique puisque r_4 et p_4 n'ont

[83]Une algèbre est une structure d'ensemble très riche puisqu'elle contient deux lois de composition internes (addition et multiplication).

[84]Le "S" de SO(n) est pour *special*, le "O" est pour *orthogonal* (*special orthogonal group*), dont les représentations de dimension N sont formées par les matrices $N \times N$ orthogonales de déterminant égal à +1 ; ces matrices sont réelles. On définit aussi le groupe O(n), qui est l'ensemble des transformations ponctuelles – c'est-à-dire ayant des points fixes – ne modifiant pas les longueurs ; ce groupe est le produit direct de SO(n) avec le groupe Z_2, formé de l'identité et de l'inversion (parité), d'où par exemple l'écriture : O(3) = SO(3) × Z_2. Une translation n'est pas une transformation ponctuelle.

pas le sens physique d'une coordonnée et d'un moment conjugué. Il s'agit cependant bel et bien d'une symétrie au sens général, symétrie induisant la dégénérescence sur le nombre l.

Il est utile de compléter cette explication un peu formelle en reprenant une partie de la belle analyse de Bander et Itzykson [47]. Dans toute la suite, on se place dans un certain sous-espace dégénéré d'énergie E, $\mathcal{E}_E$; en vertu des relations de commutation de $\vec{R}$ et $\vec{L}$, ces deux opérateurs ont tous leurs éléments de matrice nuls entre ce sous-espace et un autre sous-espace dégénéré $\mathcal{E}_{E'}$ d'énergie $E' \neq E$, ce qui autorise bien, tant que seuls ces opérateurs sont pertinents, à raisonner exclusivement dans $\mathcal{E}_E$. En ce sens, ce sous-espace est isolé, complètement indépendant de tous les autres $\mathcal{E}_{E' \neq E}$.

Compte tenu de la définition (19.247), on a maintenant :

$$\vec{L}^{\,2} + \vec{R}'^{\,2} + \hbar^2 = -\frac{\mu e'^{\,4}}{2E} \ , \qquad \vec{R}'.\vec{L} = 0 \ ; \tag{19.253}$$

par ailleurs, on vérifie sans peine que :

$$[L_u \pm R'_u,\, L_v \pm R'_v] = 2\mathrm{i}\hbar\varepsilon_{uvw}(L_w \pm R'_w) \ , \tag{19.254}$$

égalités qui prouvent que chacun des deux vecteurs $\vec{\mathcal{J}}_\pm \overset{\text{déf}}{=} \frac{1}{2}(\vec{L} \pm \vec{R}')$ satisfait les relations *caractéristiques* d'un moment cinétique, et se révèle ainsi être un authentique moment cinétique (abstrait). De la théorie du moment cinétique (chapitre 18, section 18.2), on peut dès lors affirmer[85] que les valeurs propres de $\vec{\mathcal{J}}_\pm^2$ sont de la forme $\hbar^2 \iota_\pm(\iota_\pm + 1)$; mais comme $\vec{R}'.\vec{L} = \vec{L}.\vec{R}' = 0$, on a $\iota_+ = \iota_- \equiv \iota$, le nombre 2ι étant forcément un entier (on ne peut pas trancher à ce stade si ι est restreint à des valeurs entières). Dans un sous-espace $\mathcal{E}_E$, $\vec{\mathcal{J}}_\pm^2$ est un opérateur scalaire, égal à $\iota(\iota+1)\mathbf{1}_{\mathcal{E}_E}$. En utilisant maintenant $\vec{R}'^{\,2} + \vec{L}^{\,2} = 2(\vec{\mathcal{J}}_+^2 + \vec{\mathcal{J}}_-^2)$, dont la valeur propre est $2\hbar^2[\iota(\iota+1) + \iota(\iota+1)]$, de (19.253) on déduit :

$$-\frac{\mu e'^{\,4}}{2E} = 4\hbar^2\iota(\iota+1) + \hbar^2 \equiv \hbar^2(2\iota+1)^2 \ , \tag{19.255}$$

soit :

$$\boxed{E = -\frac{\mu e'^{\,4}}{2(2\iota+1)^2\hbar^2}} \tag{19.256}$$

On reconnaît l'expression de l'énergie (19.156), à condition d'identifier $2\iota + 1$ et n ; cela étant fait, et comme $n \in \mathbb{N}^*$, on voit que le nombre ι peut prendre toutes les valeurs entières *et* demi-entières $0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$. Écrivant enfin $\vec{L} = \frac{1}{2}(\vec{\mathcal{J}}_+ + \vec{\mathcal{J}}_-)$, le théorème d'addition de deux moments cinétiques montre que pour un ι donné (donc un n donné), les valeurs possibles de l sont $0, 1, 2, \ldots, 2\iota \equiv n - 1$, donnant une dégénérescence totale $\sum_{l=0}^{2\iota}(2l+1) = n^2$.

Une autre méthode est due à Fock [45], qui écrit l'équation aux valeurs propres :

$$\left(\frac{\vec{p}^{\,2}}{2\mu} - \frac{e'^2}{r}\right)\psi(\vec{r}) = E\psi(\vec{r}) \tag{19.257}$$

[85] On se souvient que la forme des valeurs propres d'un moment cinétique est une conséquence directe des seules relations de commutation caractéristiques (18.3).

en représentation-p. En introduisant la transformée de Fourier (régularisée[86] par un facteur exponentiel) du potentiel Coulombien, le théorème de convolution permet d'écrire l'équation suivante pour la fonction propre $\phi(\vec{p})$ en représentation-p :

$$\Big(\frac{\vec{p}^{\,2}}{2\mu} - E\Big)\phi(\vec{p}) = \frac{e'^{\,2}}{2\pi^2\hbar}\int \frac{1}{\|\vec{p}\,' - \vec{p}\,\|^2}\,\phi(\vec{p}\,')\,\mathrm{d}^3p' \ . \tag{19.259}$$

Comme on va le voir, il est utile de récrire cette équation en dimension entière quelconque D dans $\mathbb{R}^D$. Dans cette situation, on a (toujours après régularisation) :

$$\frac{1}{r} = \frac{1}{\pi\Omega(D)}\int_{\mathbb{R}^D}\frac{1}{\|\vec{k}\,\|^{D-1}}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\,\mathrm{d}^Dk \ , \tag{19.260}$$

où[87] $\Omega(D)$ est la surface de la sphère de rayon unité (angle solide) dans $\mathbb{R}^D$. En posant $E = -\frac{p_0^2}{2\mu}$, l'équation (19.259) se généralise comme suit dans $\mathbb{R}^D$:

$$(\vec{p}^{\,2} + p_0^2)\phi(\vec{p}) = \frac{\mu e'^{\,2}}{\hbar}\frac{\Gamma(\frac{D-1}{2})}{\pi^{(D+1)/2}}\int_{\mathbb{R}^D}\frac{1}{\|\vec{p}\,' - \vec{p}\,\|^{D-1}}\,\phi(\vec{p}\,')\,\mathrm{d}^Dp' \ . \tag{19.261}$$

À ce stade, l'équation ne semble présenter que l'invariance par rotation dans $\mathbb{R}^D$, rien de plus. En réalité, l'invariance est encore respectée dans un espace de dimension $D+1$. L'idée de Fock est d'utiliser une transformation stéréographique $\mathcal{S}$ associant un point de $\mathbb{R}^D$ à un point de $\mathbb{R}^{D+1}$; géométriquement, pour $D = 2$, la transformation se définit comme suit. On considère la sphère unité dans $\mathbb{R}^{2+1}$, et son plan équatorial isomorphe à $\mathbb{R}^2$. À un point quelconque du plan $\mathbb{R}^2$, défini par le vecteur $\vec{p}/p_0$, on associe le point de la sphère (dans $\mathbb{R}^3$) qui est l'intersection de celle-ci avec le segment joignant l'extrémité de $\vec{p}/p_0$ au pôle nord de la sphère[88] ; le point de la sphère est l'extrémité d'un vecteur $\vec{\varpi}$ de $\mathbb{R}^3$. Tous les points du plan ayant $\|\vec{p}\,\| > p_0$ sont envoyés dans l'hémisphère nord, ceux ayant ayant $\|\vec{p}\,\| < p_0$ dans l'hémisphère sud. Cette transformation se définit de façon analogue dans $\mathbb{R}^D$; il n'est pas difficile de voir que si $\vec{n}$ est le vecteur (unitaire) associé au pôle nord, le vecteur $\vec{\varpi}$ fixant l'image de $\vec{p}$ est :

$$\vec{p} \overset{\mathcal{S}}{\rightarrow} \vec{\varpi} \overset{\text{déf}}{=} \mathcal{S}(\vec{p}) \quad\Longleftrightarrow\quad \vec{\varpi} = \frac{p^2 - p_0^2}{p^2 + p_0^2}\,\vec{n} + \frac{2p_0}{p^2 + p_0^2}\,\vec{p} \ . \tag{19.262}$$

On en déduit d'abord :

$$\|\vec{p}\,' - \vec{p}\,\|^2 = \frac{1}{4p_0^2}(p^2 + p_0^2)(p'^2 + p_0^2)\|\vec{\varpi}\,' - \vec{\varpi}\,\|^2 \ ; \tag{19.263}$$

---

[86] On définit précisément la transformée de Fourier de $\frac{1}{r}$ comme :

$$\mathcal{F}\Big[\frac{1}{r}\Big] \overset{\text{déf}}{=} \lim_{\lambda\rightarrow 0_+}\int_{\mathbb{R}^3}\frac{\mathrm{e}^{-\lambda r}}{r}\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\mathrm{d}^3r = \lim_{\lambda\rightarrow 0_+}\frac{4\pi}{k^2+\lambda^2} = \frac{4\pi}{k^2} \ . \tag{19.258}$$

[87] Suivant un résultat classique [54], $\Omega(D) = \frac{2\pi^{D/2}}{\Gamma(D/2)}$; Γ est la fonction d'Euler de deuxième espèce (voir note 43, p. 805).

[88] C'est exactement la même transformation qui permet de définir le point à l'infini du plan complexe $\mathbb{C}$, la sphère étant appelée dans ce contexte *sphère de Riemann* ([55], § 1.5). Le point à l'infini de $\mathbb{C}$ est l'antécédent par $\mathcal{S}$ du pôle nord.

par ailleurs, on a :

$$\int_{\mathbb{R}^D} f(\vec{p}) \left(\frac{2p_0}{p^2+p_0^2}\right)^D \mathrm{d}^D p = \int_{\text{sphère} \subset \mathbb{R}^{D+1}} F(\vec{\varpi}\,) \,\mathrm{d}^{D+1}\varpi \tag{19.264}$$

où $F(\vec{\varpi}) \stackrel{\text{déf}}{=} f(\vec{p} = \mathcal{S}^{-1}\vec{\varpi})$. Posons maintenant $\Phi(\vec{\varpi}) \stackrel{\text{déf}}{=} p_0^{-1/2}\left(\frac{p^2+p_0^2}{2p_0}\right)^{(D+1)/2} \phi(\vec{p})$; le report de toutes ces expressions dans l'équation intégrale (19.261) conduit à :

$$\Phi(\vec{\varpi}) = \frac{\mu e'^2}{2p_0\hbar}\frac{\Gamma(\frac{D-1}{2})}{\pi^{(D+1)/2}} \int_{\text{sphère} \subset \mathbb{R}^{D+1}} \frac{1}{\|\vec{\varpi}\,' - \vec{\varpi}\,\|^{D-1}}\, \Phi(\vec{\varpi}\,') \,\mathrm{d}^{D+1}\varpi\,' \ . \tag{19.265}$$

Cette équation montre que le problème possède la symétrie sphérique dans l'espace de dimension $D+1$. Revenant à $D = 3$, ceci signifie qu'il y a invariance par rapport au groupe SO(4) des rotations dans $\mathbb{R}^4$, l'équation (19.265) étant :

$$\Phi(\vec{\varpi}) = \frac{\mu}{m_e}\frac{\hbar}{4\pi^{\frac{3}{2}} a_0 p_0} \int_{\text{sphère} \subset \mathbb{R}^4} \frac{1}{\|\vec{\varpi}\,' - \vec{\varpi}\,\|^2}\, \Phi(\vec{\varpi}\,') \,\mathrm{d}^4\varpi\,' \ . \tag{19.266}$$

Il y a donc une correspondance biunivoque entre les fonctions propres $\phi(\vec{p})$ satisfaisant l'équation aux valeurs propres (19.259), et les fonctions $\Phi(\vec{\varpi}\,)$ satisfaisant (19.266) ; cette transformation est unitaire [47], et permet ainsi de relier les fonctions propres de carré sommable dans $\mathbb{R}^3$ (associées aux états liés du potentiel Coulombien) et les fonctions $\Phi(\vec{\varpi}\,)$ définies sur la sphère dans $\mathbb{R}^4$, solutions de (19.266), également de carré sommable puisque la transformation conserve tout produit scalaire.

L'existence de la symétrie dynamique du champ coulombien est une propriété remarquable, dont la conséquence certainement la plus spectaculaire est l'indépendance de l'énergie par rapport au nombre l attaché à la valeur du module du moment cinétique. Clairement, l'existence d'une grande symétrie contraint fortement les degrés de liberté d'un système en *verrouillant* sa dynamique ; il est d'ailleurs loisible de se demander si la coïncidence énigmatique entre les formules de Bohr et de Schrödinger (voir note 52 p. 809) n'est pas une autre conséquence de ce verrouillage. Aucune discussion établissant un lien précis entre ces deux spécificités ne semble disponible dans la littérature.

19.3 Exercices et problèmes

19.3.1 Démonstration de l'égalité (19.21)

1. Trouver l'expression de p_r dans $\mathbb{R}^D$.

2. Montrer que $[r,\, p_r] = \mathrm{i}\hbar\mathbf{1}$ quelle que soit la dimension D.

3. Montrer que, dans $\mathbb{R}^D$, p_r^2 est donné par l'expression (19.21).

19.3.2 Champ central dans le plan

Une particule de masse μ se déplace dans le plan $\mathbb{R}^2$ dans le potentiel central $V(r)$.

1. Soit (r, θ) les coordonnées polaires ; écrire l'expression de $\vec{p}^2$ en fonction de r et θ.
2. En déduire l'équation propre pour $\psi(\vec{r})$; montrer que les variables se séparent.
3. On pose $\psi(\vec{r}) = r^{-1/2}u(r)\,\mathrm{e}^{\mathrm{i}m\theta}$.
   (a) Que peut-on dire du nombre m ?
   (b) Écrire l'équation différentielle satisfaite par la fonction $u(r)$.

   En utilisant les résultats de ce chapitre et en faisant le moins de calculs possible, traiter les cas suivants :

   (a) Particule libre.
   (b) Puits "carré" circulaire[89] de rayon a, défini comme en (19.93).
   Examiner en particulier les états de moment cinétique nul et discuter le nombre d'états liés en fonction de $k_0 \stackrel{\text{déf}}{=} \hbar^{-1}\sqrt{2\mu V_0}$.
   (c) Champ Coulombien[90] $V(r) = -\frac{e'^2}{r}$.

19.3.3 Difficultés du puits δ en dimension 2

Lorsque l'on est en présence d'un puits de potentiel très profond et de faible extension spatiale, on est tenté de le modéliser par une fonction de Dirac. Cette assimilation ne présente aucune difficulté en dimension $d = 1$. Dans la suite, on met en évidence la pathologie d'un tel traitement en dimension supérieure.

Dans la suite, on ne considère que les états liés. Pour fixer les idées, on se place d'abord à deux dimensions. Une représentation de la fonction de Dirac dans $\mathbb{R}^2$ est :

$$\delta(\vec{r}) = \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} \mathrm{d}^2k\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \ . \tag{19.268}$$

Le puits étroit et profond est modélisé par le potentiel $V = -\lambda\delta(\vec{r})$, où λ est une constante réelle (quelle est sa dimension ?).

---

[89] L'équation :

$$Z''(\rho) + \frac{1}{\rho}Z'(\rho) - \left(1 + \frac{\nu^2}{\rho^2}\right)Z(\rho) = 0 \tag{19.267}$$

possède une solution décroissant exponentiellement à l'infini ([38], § 8.494.1), notée $H_\nu^{(1)}(\mathrm{i}\rho)$, proportionnelle à la fonction de Bessel notée traditionnellement $K_\nu(\rho)$; pour $\nu = \frac{1}{2}$, on a $K_{1/2}(\rho) = \sqrt{\frac{\pi}{2z}}\,\mathrm{e}^{-\rho}$. Par ailleurs, $J_{1/2}(\rho) = \sqrt{\frac{2}{\pi\rho}}\,\sin\rho$.

[90] Cette appellation – usuelle – est en fait impropre, ou plutôt dépend de ce que l'on appelle *champ Coulombien*. Une première acception consiste à retenir l'idée d'une force en r^{-2}, qui donne effectivement $V(r) \propto r^{-1}$; une autre possibilité est de se référer à l'équation de Poisson $\Delta V = 0$ (en dehors de la source ponctuelle), soit $V''(r) + \frac{1}{r}V'(r) = 0$. Une première intégration donne $V'(r) = \frac{C}{r}$, d'où $V(r) = C\ln\frac{r}{r_0}$.

1. Écrire l'équation donnant les valeurs et fonctions propres E et $\psi(\vec{r})$. Quel est le signe de E ?

2. On pose :
$$E = -\frac{\hbar^2 K^2}{2m} \ , K > 0 \ , \qquad \eta = \frac{2m\lambda}{\hbar^2} \ . \tag{19.269}$$
Quelle est la dimension de η ? Récrire l'équation obtenue en **1.** à l'aide de K et η, et en y faisant apparaître $\psi(\vec{0})$.

3. $\phi(\vec{k})$ désignant la transformée de Fourier de $\psi(\vec{r})$:
$$\phi(\vec{k}) = \int_{\mathbb{R}^2} \mathrm{d}^2 r \, \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}} \, \psi(\vec{r}) \quad \Longleftrightarrow \quad \psi(\vec{r}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \mathrm{d}^2 k \, \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \, \phi(\vec{k}) \ , \tag{19.270}$$
écrire l'équation satisfaite par la fonction $\phi(\vec{k})$.

4. En déduire l'expression de $\phi(\vec{k})$ sous la forme d'une fraction rationnelle.

5. Utiliser la définition (19.270) pour écrire une expression intégrale de $\psi(\vec{0})$, et montrer que $1/\eta$ est formellement donné par une certaine intégrale, notée I.

6. Cette dernière est-elle bien définie ? Conclusion ?

7. Pour une dimension d'espace quelconque, D, les diverses représentations intégrales rencontrées ci-dessus se généralisent aisément. Par exemple :
$$\delta(\vec{r}) = \frac{1}{(2\pi)^D} \int_{\mathbb{R}^D} \mathrm{d}^D k \, \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \ , \qquad \mathrm{d}^D k = C(D) \, k^{D-1} \mathrm{d}k \tag{19.271}$$
$C(D)$ étant une fonction connue de D (qui vaut 2π pour $D = 2$, 4π pour $D = 3$, etc.). Trouver la dimension D_c au-dessus de laquelle la modélisation élémentaire par un puits de Dirac soulève une difficulté.

19.3.4 La coquille de Dirac

En utilisant les résultats du problème 15.6.8 du Tome I, étudier les états liés sphériques du potentiel $V(r) = -g\delta(r - a)$, $g > 0$, $a > 0$. Le puits $-g\delta(r)$ a-t-il un état lié ?

19.3.5 Puits "carré" circulaire et limite δ

Une particule chargée (masse m, charge q) se déplace dans le plan et a pour énergie potentielle :
$$V(r) = \begin{cases} -V_0 & \text{si } r < a \\ 0 & \text{si } r > a \end{cases} \tag{19.272}$$
V_0 étant positif. Dans toute la suite, on ne considère que les états liés sphériques d'énergie E, $R(r)$ désignant la fonction radiale. On pose :
$$k = \hbar^{-1}\sqrt{2m|E|} \ , \qquad k_0 = \hbar^{-1}\sqrt{2m(V_0 - |E|)} \tag{19.273}$$

1. Écrire les équations pour $R(r)$ dans les deux régions $r < a$ et $r > a$.

2. Montrer que les solutions sont les fonctions de Bessel J_0 et K_0 (voir note 89).

3. Sachant que $J_0'(z) = -J_1(z)$ et $K_0'(z) = -K_1(z)$, exprimer les conditions de raccordement à l'aide des rapports $\frac{J_1(k_0 a)}{J_0(k_0 a)}$ et $\frac{K_1(ka)}{K_0(ka)}$.

4. On analyse ici la limite $V_0 \to +\infty$, $a \to 0$, avec la contrainte $V_0 a^2 = \mathrm{C^{ste}} \equiv g$, de sorte que le potentiel tende vers $-g\delta_{\mathbb{R}^2}(\vec{r})$. Il s'agit de montrer que la condition trouvée en 3 ne peut être satisfaite à la limite dans l'hypothèse où la valeur propre E reste *finie*.

   (a) Trouver le comportement du membre impliquant J_0 et J_1.

   (b) Sachant[91] que pour $|z| \ll 1$, $K_0(z) \simeq -\ln\frac{z}{2} - \mathrm{C}$, $K_1(z) \simeq \frac{1}{z} + \frac{z}{2}\ln\frac{z}{2}$, montrer que l'équation de raccordement ne peut pas être satisfaite avec l'hypothèse que la valeur propre E est finie à la limite.

5. Commenter ces résultats.

19.3.6 Particule libre en coordonnées sphériques

La résolution en coordonnées sphériques du problème de la particule libre (masse μ) introduit les fonctions propres $\psi_{klm}(r,\theta,\phi) = C_{kl}\, j_l(kr) Y_{lm}(\theta,\phi)$ où les j_l désignent toujours les fonctions de Bessel sphériques. k est le nombre quantique relié à l'énergie E par $k = \sqrt{2\mu E}/\hbar$ et varie continûment entre 0 et $+\infty$. Toute "bonne" fonction $\Psi(\vec{r})$ peut être décomposée suivant :

$$\Psi(\vec{r}) = \int_0^{+\infty} \mathrm{d}k \sum_{l=0}^{+\infty} \sum_{m=-l}^{+l} f_{lm}(k)\, j_l(kr)\, Y_{lm}(\theta,\phi). \tag{19.274}$$

1. Comment se simplifie un tel développement dans le cas d'une fonction Ψ à symétrie sphérique ?

2. Normaliser la partie radiale $R_{k0}(r) \equiv C_{k0}\, j_0(kr)$ de sorte que :

$$\int_0^{+\infty} R_{k0}(r)\, R_{k'0}(r)\, r^2 \mathrm{d}r = \delta(k-k'). \tag{19.275}$$

3. Soit $\Psi(r)$ une fonction à symétrie sphérique admettant le développement :

$$\Psi(r) = \int_0^{+\infty} f(k)\, R_{k0}(r)\, \mathrm{d}k \ ; \tag{19.276}$$

   exprimer la fonction $f(k)$ sous la forme d'une intégrale.

[91] On adopte ici les notations de Gradsteyn et Rhyzhik [38], où $\mathrm{C} = 0,577\,215\,665...$; dans certains ouvrages, la constante C est notée γ.

4. Appliquer ces résultats au cas de la gaussienne $\Psi(r) = C\mathrm{e}^{-\alpha r^2/2}$, après l'avoir normalisée. Trouver la fonction $f(k)$ et en déduire la densité de probabilité $P(E)$ des résultats de mesure de l'énergie sur un paquet d'ondes gaussien[92].

Rappels

$$\int_0^{+\infty} x^{z-1}\mathrm{e}^{-x}\,\mathrm{d}x = \Gamma(z) \quad (\Re z > 0)\ , \quad \Gamma(z+1) = z\Gamma(z)\ , \quad \Gamma(1/2) = \sqrt{\pi}. \quad (19.277)$$

19.3.7 Puits sphérique infini

Une particule de masse μ est confinée dans un puits sphérique infiniment profond :

$$V(r) = \begin{cases} 0 & \text{si } r < a \\ +\infty & \text{si } r > a \end{cases} \quad (19.278)$$

et on pose $\psi(r, \theta, \phi) = r^{-1}\,u(r)\,Y_{lm}(\theta,\,\phi)$.

1. Écrire l'équation donnant la fonction radiale $u(r)$.
2. On pose $E = \frac{\hbar^2 k^2}{2\mu}$, $\rho = kr$, $u(r) = \rho^{1/2}\,v(\rho)$. Écrire l'équation satisfaite par $v(\rho)$.
3. En déduire que $v(\rho)$ est une combinaison linéaire de deux fonctions de Bessel ; expliquer pourquoi il faut rejeter la solution singulière à l'origine.
4. Exprimer la fonction radiale $R(r)$ à l'aide des fonctions de Bessel sphériques j_l :

$$j_l(x) \overset{\text{déf}}{=} \sqrt{\frac{\pi}{2x}}\,J_{l+\frac{1}{2}}(x) \quad (19.279)$$

5. Écrire explicitement les équations donnant les valeurs possibles de k pour les états s, p et d. Tracer le schéma des premiers niveaux d'énergie.

Quelques résultats à propos des fonctions de Bessel

1. La fonction de Bessel J_ν est solution de l'équation de Bessel :

$$J_\nu''(x) + \frac{1}{x}J_\nu'(x) + \left(1 - \frac{\nu^2}{x^2}\right)J_\nu = 0\ . \quad (19.280)$$

[92]Poser $E_0 = \frac{\alpha\hbar^2}{2\mu}$

2. J_ν a un développement du genre :

$$J_\nu(x) = x^\nu \sum_{k=0}^{+\infty} c_k \, x^k \tag{19.281}$$

où les coefficients c_k sont connus mais sans intérêt ici (le développement est donné en (19.76)).

3. les premières fonctions de Bessel sphériques sont données en (19.83).

4. $\forall \nu \in \mathbb{R}$, J_ν a une infinité de zéros réels.

19.3.8 Désintégration du tritium

Le tritium est un isotope de l'hydrogène, noté ^{3}H. Son noyau est radioactif et se transforme par désintégration β en un noyau d'hélium ^{3}He. La réaction élémentaire est :

$$^3\mathrm{H} \quad \longrightarrow \quad ^3\mathrm{He}^+ + e^- + \bar{\nu} \tag{19.282}$$

où $\bar{\nu}$ désigne un antineutrino, qui ne joue aucun rôle dans la suite. L'électron issu du noyau a une énergie élevée (de l'ordre de 15 keV), donc une vitesse élevée : ainsi, il est légitime de considérer qu'il sort du système instantanément. En conséquence, dans l'atome ionisé He$^+$, la fonction d'onde de l'unique électron est initialement quasiment indiscernable de celle décrivant l'état fondamental du Tritium[93].

1. Pour cet électron atomique, écrire son Hamiltonien H_i avant la désintégration, et son Hamiltonien H_f après la désintégration.

2. Quelles sont les énergies E'_n et les fonctions propres ϕ_{nlm} de He$^+$?

3. Quelle est l'énergie E_f de l'électron atomique juste après la désintégration (l'exprimer en fonction de E_I) ?

4. Donner l'expression de P_{nlm}, probabilité de trouver l'électron atomique dans l'état propre ϕ_{nlm} de l'ion He$^+$. Combien vaut P_{nlm} si $(l,\, m) \neq (0,\, 0)$?

5. Dans la suite, on note simplement $P_{n00} \equiv P_n$. Calculer explicitement P_1.

6. Un calcul numérique donne $\sum_{n=2}^{+\infty} P_n \simeq 0,2714$. En déduire la probabilité p de trouver l'électron dans l'un quelconque des états liés de He$^+$.

7. De la même façon, on peut montrer que $\sum_{n=2}^{+\infty} n^{-2} P_n \simeq 0,0643$. En déduire la valeur moyenne $\langle E \rangle \stackrel{\text{déf}}{=} \sum_{n=1}^{+\infty} P_n \, E'_n$ en fonction de E_I.

8. Interpréter physiquement les inégalités $\langle E \rangle \neq E_\mathrm{f}$ et $p < 1$.

[93] La désintégration est si rapide à l'échelle atomique que la fonction d'onde électronique n'a pas le temps de changer.

19.3.9 États liés sphériques du deutéron

Le deutéron est constitué d'une paire liée proton - neutron ; leur attraction à courte portée $V(r)$ est modélisée suivant $V(r) = -V_0 \mathrm{e}^{-r/a}$, $V_0 > 0$, $a > 0$.

1. Dessiner le potentiel en fonction de r.
2. Soit μ la masse réduite ; en confondant les masses du proton et du neutron, notée chacune M, combien vaut μ ? Dans la suite, on pose $k_0 \stackrel{\text{déf}}{=} 8\mu V_0/\hbar^2$.
3. On se borne aux états liés de type s, d'énergie $E \stackrel{\text{déf}}{=} -\frac{\hbar^2 k^2}{8\mu}$, $k > 0$. Soit $R(r)$ la fonction radiale et $u(r) \stackrel{\text{déf}}{=} rR(r)$. Écrire l'équation satisfaite par $u(r)$.
4. On introduit la nouvelle variable $X \stackrel{\text{déf}}{=} ak_0\mathrm{e}^{-\frac{r}{2a}}$ et la fonction $w(X) \stackrel{\text{déf}}{=} u(r)$. Écrire l'équation différentielle satisfaite par $w(X)$.
5. En déduire que $w(X)$ est une combinaison linéaire de deux fonctions de Bessel[94] :
$$w(X) = A\,J_\lambda(X) + B\,J_{-\lambda}(X) \ , \tag{19.283}$$
et identifier le paramètre $\lambda > 0$.
6. Que vaut la constante B ?
7. À l'aide de la bonne condition aux limites, écrire – exprimée à l'aide d'une fonction de Bessel – l'équation donnant les valeurs possibles de l'énergie E.
8. Préciser la condition sur la profondeur V_0 du puits pour qu'il existe au moins un état lié.

19.3.10 Oscillateur harmonique à trois dimensions

1. Il s'agit d'abord de trouver, pour l'oscillateur harmonique isotrope (μ, ω), les fonctions propres communes à $(H,\ \vec{L}^2,\ L_z)$. On pose $\psi(r,\theta,\phi) = r^{-1}u(r)Y_{lm}(\theta,\phi)$.
   (a) Écrire l'équation satisfaite par $u(r)$.
   (b) Trouver la constante α en supposant que $u \sim \mathrm{e}^{-\alpha r^2/2}$ quand $r \to \infty$.
   (c) Compte tenu de ce qui précède, on pose $u(r) = \mathrm{e}^{-\alpha r^2/2}\,v(r)$. Trouver l'équation différentielle pour la nouvelle fonction inconnue $v(r)$.
   (d) Chercher v sous la forme d'une série entière $v(r) = r^s \sum_{m=0}^{+\infty} c_m r^m$, $c_0 \neq 0$ et trouver la relation de récurrence entre les coefficients c_m. En examinant les deux premières équations, pour c_0 et c_1, montrer que $v(r)$ est de la forme :
$$v(r) = r^{l+1} \sum_{p=0}^{+\infty} c_{2p}\, r^{2p}. \tag{19.284}$$

[94] Voir les éléments rappelés à la fin de l'énoncé du problème 19.3.7.

(e) En trouvant une série de comparaison, montrer que la série apparaissant dans $v(r)$, éq. (19.284), se comporte comme $e^{+\alpha r^2}$ pour $\alpha r^2 \gg 1$. En déduire qu'il doit exister un entier $p_0 \geq 0$ tel que :

$$k^2 - \alpha(4p_0 + 2l + 3) = 0. \qquad (19.285)$$

(f) On pose $n = 2p_0 + l$. Trouver les valeurs de l'énergie, E_n, et préciser la dégénérescence des niveaux.

(g) Comment se compare la partie radiale des états s à la fonction propre d'un oscillateur à une dimension ?

2. On se place maintenant en coordonnées cartésiennes.

(a) Expliquer pourquoi l'on peut chercher les états propres sous la forme

$$\psi(x, y, z) = \psi_x(x)\psi_y(y)\psi_z(z) \qquad (19.286)$$

(b) Que sont les fonctions propres $\psi_u(u)$?

(c) Trouver sans calcul l'expression de l'énergie $E_{n_x n_y n_z}$. Quelle est la dégénérescence ?

(d) Supposant maintenant l'oscillateur *anisotrope*, la pulsation le long de Oz étant $\tilde{\omega} \neq \omega$, trouver les nouvelles énergies $\tilde{E}_{n_x n_y n_z}$ et la dégénérescence des états. Examiner les limites $\tilde{\omega} \to 0$ et $\tilde{\omega} \to \infty$ et les interpréter physiquement.

19.3.11 Sur l'atome d'hydrogène

1. Il s'agit de trouver le champ électrostatique moyen en un point $\vec{R}$ dans l'atome d'hydrogène à l'état fondamental.

(a) Écrire l'expression du champ élémentaire en $\vec{R}$ créé par l'électron situé en $\vec{r}$. Trouver le champ électronique moyen.

(b) En déduire le champ électrostatique total au point $\vec{R}$; tracer son module en fonction de R. Commenter.

(c) Calculer le potentiel électrostatique $W(R)$.

(d) Quelle est la *self-energy* de l'électron ?

2. Soit l'atome d'hydrogène dans un état propre commun à H et L^2, $|n, l\rangle$.

(a) Si on traite de façon classique un atome de même énergie E_n et de même moment cinétique L, quelles sont les valeurs minimale et maximale de la distance r entre le proton et l'électron ? On exprimera ces distances en fonction du rayon de Bohr a_0 et des nombres quantiques n et l. À quelle condition la distance minimale peut-elle être nulle ?

(b) Dans l'état fondamental, quelle est la probabilité pour que la distance entre le proton et l'électron soit supérieure à la distance maximale permise par la mécanique classique ?

3. La densité radiale est définie comme :

$$\rho_{nl}(r) \stackrel{\text{déf}}{=} r^2 \left[R_{nl}(r)\right]^2 \ , \tag{19.287}$$

où $R_{nl}(r)$ est la fonction radiale normalisée.

(a) Pour n donné, écrire l'expression de la densité radiale dans l'état de plus grand moment cinétique.

(b) Trouver la valeur r_{max} où cette densité est maximum, et trouver l'expression de la valeur du maximum.

(c) En utilisant la formule de Stirling $N! \simeq \sqrt{2\pi N}\left(\frac{N}{\text{e}}\right)^N$ ($N \gg 1$), montrer que ce maximum décroît comme $n^{-3/2}$ pour $n \gg 1$.

4. Quel est l'écart quadratique de p^2 dans l'état 1s ?

5. Dans la suite, $\langle r^s \rangle$ désigne la valeur moyenne de r^s dans un état propre $\psi_{nlm}(r,\,\theta,\phi)$.

(a) En manipulant l'équation aux valeurs propres pour $u(r) = rR(r)$, démontrer la relation de Kramers :

$$\frac{s+1}{n^2}\langle r^s\rangle - (2s+1)a_0\langle r^{s-1}\rangle + \frac{s}{4}[(2l+1)^2 - s^2]a_0^2\langle r^{s-2}\rangle = 0 \quad (s > -2l-1) \ . \tag{19.288}$$

(b) En déduire $\langle r\rangle$, $\langle r^2\rangle$, $\langle r^{-2}\rangle$ et l'écart-type relatif $\Delta r/\langle r\rangle$.

(c) Examiner les deux limites ($l = 0, n \to +\infty$), ($l = n-1, n \to +\infty$) et en déduire que la limite classique n'est obtenue que si *tous* les nombres quantiques tendent vers l'infini.

6. Trouver Δp_r dans un état de type s, et en déduire le produit $\Delta p_r \Delta r$.

19.3.12 Compléments sur les fonctions radiales hydrogénoïdes

Il s'agit de compléter les résultats techniques donnés dans l'aparté p. 820.

1. P_l désignant un polynôme de Legendre et j_l une fonction de Bessel sphérique, on sait que (voir (19.91)) :

$$\text{e}^{\,ikr\cos\theta} = \sum_{l=0}^{+\infty} (2l+1)\,\text{i}^l\, j_l(kr)\, P_l(\cos\theta) \ . \tag{19.289}$$

(a) Utiliser le théorème d'addition des harmoniques sphériques, (18.121), pour démontrer :

$$e^{i\vec{k}.\vec{r}} = 4\pi \sum_{l=0}^{+\infty} \sum_{m=-l}^{+l} i^l \, j_l(kr) \, Y_{lm}^*(\theta, \phi) \, Y_{lm}(\Theta, \Phi) \ . \tag{19.290}$$

où (θ, ϕ) est la direction de $\vec{r}$ et (Θ, Φ) la direction de $\vec{k}$.

(b) Soit $\psi_{nlm}(\vec{r})$ et $\phi_{nlm}(\vec{p})$ les fonctions propres en représentation-q et -p :

$$\phi_{nlm}(\vec{p}) = (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} e^{\frac{1}{i\hbar}\vec{p}.\vec{r}} \, \psi_{nlm}(\vec{r}) \, d^3r \ . \tag{19.291}$$

Montrer que $\phi_{nlm}(\vec{p})$ est de la forme $\chi_{nl}(p)Y_{lm}(\Theta, \Phi)$, où (Θ, Φ) est la direction de $\vec{p}$; exprimer χ_{nl} en fonction de $R_{nl}(r)$.

(c) Trouver χ_{10}, χ_{20} et χ_{21} connaissant R_{10}, R_{20} et R_{21} (voir (19.187)).

2. En raisonnant directement sur l'équation de Laplace (19.137), démontrer l'égalité exprimant la transformation de Kummer (voir p. 810).

3. Les polynômes de Laguerre, $L_p(z)$, sont solutions[95] de l'équation :

$$z\,L''(z) + (1-z)\,L'(z) + p\,L(z) = 0 \qquad (p \text{ entier}) \ . \tag{19.292}$$

(a) Pour trouver L, on utilise la méthode de Laplace en posant :

$$L(z) = \int_C e^{-zt} \, f(t) \, dt \ . \tag{19.293}$$

Le contour C est à préciser dans la suite. Écrire la représentation intégrale des dérivées L' et L''.

(b) Exprimer zL'' et zL' à l'aide d'une intégrale contenant f', en faisant dès à présent l'hypothèse que le contour C est choisi de sorte que les termes tout intégrés sont nuls. En déduire une équation différentielle du premier ordre pour la fonction $f(t)$.

(c) Intégrer cette équation et en déduire que $L_p(z)$ peut s'écrire :

$$L_p(z) = A \int_C \frac{e^{-zt}}{t^{p+1}} \, (1+t)^p \, dt \ . \tag{19.294}$$

Trouver la constante A sachant que la convention habituelle est de poser $L_p(0) = p!$.

(d) Déduire de c) que $L_p(z)$ est donné par :

$$L_p(z) = e^z \frac{d^p}{dz^p} \, e^{-z} \, z^p \ ; \tag{19.295}$$

écrire l'expression explicite de $L_p(z)$.

[95] Ce ne sont pas les seules.

(e) Trouver la fonction génératrice des L_p :

$$F(s, z) = \sum_{p=0}^{+\infty} \frac{L_p(z)}{p!}\, s^p \qquad (|s| \leq 1) \ . \tag{19.296}$$

(f) Les polynômes associés de Laguerre, $L_p^q(z)$ sont définis comme suit :

$$L_p^q(z) \overset{\text{déf}}{=} (-1)^q \frac{\mathrm{d}^q}{\mathrm{d}z^q} L_{p+q}(z) \ . \tag{19.297}$$

Quel est leur degré ? Montrer qu'ils satisfont l'équation différentielle :

$$z\, L''(z) + (q + 1 - z)\, L'(z) + p\, L(z) = 0 \qquad (p,\, q \text{ entiers}) \ . \tag{19.298}$$

Trouver la fonction génératrice des $L_p^q(z)$:

$$F_q(s, z) \overset{\text{déf}}{=} \sum_{p=0}^{+\infty} \frac{L_p^q(z)}{(p+q)!}\, s^p \qquad (|s| \leq 1) \ . \tag{19.299}$$

(g) Soit :

$$G_q(s, t, \lambda) \overset{\text{déf}}{=} \int_0^{+\infty} \frac{\mathrm{e}^{-sz/(1-s)}\, \mathrm{e}^{-tz/(1-t)}}{(1-s)^{q+1}\,(1-t)^{q+1}}\, z^q \mathrm{e}^{-\lambda z}\, \mathrm{d}z \ . \tag{19.300}$$

Calculer $G_q(s, t, \lambda)$ directement, puis écrire son développement en puissances de s et t. Faisant $\lambda = 1$, en déduire que les $L_p^q(z)$ satisfont la relation d'orthogonalité :

$$\int_0^{+\infty} \rho^q \mathrm{e}^{-\rho}\, L_p^q(\rho)\, L_{p'}^q(\rho)\, \mathrm{d}\rho = \frac{[(p+q)!]^3}{p!}\, \delta_{pp'} \ . \tag{19.301}$$

(h) Sachant que :

$$R_{nl}(r) = C_{nl}\, (2\rho)^l\, L_{n-l-1}^{2l+1}(2\rho)\, \mathrm{e}^{-\rho} \ , \qquad \rho = k_n r, \quad k_n = \frac{1}{n a_0} \ , \tag{19.302}$$

analyser $[\partial G_q / \partial \lambda]_{\lambda=1}$ et montrer que la constante de normalisation C_{nl} vaut :

$$C_{nl} = a_0^{-3/2}\, \frac{2}{n^2} \sqrt{\frac{(n-l-1)!}{[(n+l)!]^3}} \ . \tag{19.303}$$

19.3.13 Méthode de Laplace et fonction hypergéométrique

Il s'agit de montrer par l'exemple l'efficacité de la méthode de Laplace pour résoudre certaines équations différentielles à coefficients variables. Cette technique permet de *réduire* l'ordre d'une équation ; si, ce faisant, on en arrive à l'ordre un, le problème est

formellement résolu puisque la solution de toute équation homogène du premier ordre s'obtient par une simple intégration (quadrature). Comme on va le comprendre, c'est effectivement ce qui se produit pour une équation du second ordre dont les coefficients sont au plus des fonctions *linéaires* de la variable.

L'exemple traité est celui de l'équation de Laplace (19.137), mais le lecteur transposera aisément la méthode à toute autre équation du même type. L'un des objectifs est d'obtenir une représentation intégrale de la fonction $\Phi(\alpha, \gamma, z)$ définie en (19.142), et qui constitue la solution analytique à l'origine de l'équation de Laplace.

Soit $L(z)$ la solution générale de l'équation de Laplace (19.137), que l'on écrit sous la forme :

$$L(z) = \int_C F(\xi)\, \mathrm{e}^{z\xi}\, \mathrm{d}\xi \ ; \tag{19.304}$$

$F(\xi)$ est une certaine fonction qui, une fois déterminée, contribuera à définir une représentation intégrale de la solution $L(z)$. C est un certain contour dans le plan complexe $\mathbb{C}$, dont les caractéristiques seront précisées au fur et à mesure de l'élaboration de la méthode. Clairement, la transformation intégrale (19.304) est une sorte de généralisation de la transformation de Laplace conventionnelle, le chemin C étant dans ce dernier cas le demi-axe réel positif $\mathbb{R}_+$. Dans un souci d'efficacité, on ne se préoccupera pas des subtilités liées à la convergence uniforme, le lecteur soucieux en la matière étant renvoyé aux ouvrages spécialisés pour se convaincre de la légitimité de certaines opérations effectuées dans la suite. Par ailleurs et afin de ne pas alourdir le sujet à traiter, on fera le cas échéant des hypothèses simplificatrices (par exemple supposer que γ est un entier naturel pair), qui ne nuisent d'ailleurs pas à la généralité du problème posé puisque dans tous les cas physiques rencontrés ci-dessus $\gamma = 2l + 2$ avec $l \in \mathbb{N}$.

1. En effectuant des intégrations par parties, montrer que :

$$zL'(z) = \xi F(\xi)\, \mathrm{e}^{z\xi}|_C - \int_C [\xi F(\xi)]'\, \mathrm{e}^{z\xi} \ , \tag{19.305}$$

et :

$$zL''(z) = \xi^2 F(\xi)\, \mathrm{e}^{z\xi}|_C - \int_C [\xi^2 F(\xi)]'\, \mathrm{e}^{z\xi} \ . \tag{19.306}$$

2. En déduire que l'on peut choisir $F(\xi)$ comme la solution d'une certaine équation différentielle du *premier* ordre, à condition de choisir le contour C de sorte que la variation le long de C de la fonction $\phi(\xi) \stackrel{\text{déf}}{=} \xi(1-\xi)F(\xi)\, \mathrm{e}^{z\xi}$ soit nulle.

3. Montrer que la solution de l'équation du premier ordre de la forme :

$$F(\xi) = \xi^\lambda (\xi - 1)^\mu \ , \tag{19.307}$$

où λ et μ sont des exposants à déterminer.

4. Déterminer précisément les singularités de la fonction $F(\xi)$.

5. Dans toute la suite, on considère exclusivement la branche de la fonction multiforme z^λ comme celle prenant des valeurs positives si z est un réel positif. Préciser les expressions de z^λ quand z est un réel positif x, ou quand $z = -x \pm \mathrm{i}0$, avec $x > 0$.

6. En déduire l'expression explicite de $F(\xi)$ quand $\xi = x$ avec $x > 1$, quand $\xi = x \pm \mathrm{i}0$ quand $0 < x < 1$, et quand $\xi = -x \pm \mathrm{i}0$ avec $x > 0$.

7. On suppose maintenant que γ est un entier naturel pair, positif ou nul. Montrer que, à une constante près, la fonction $L(z)$ a pour expression :

$$L(z) = \int_0^1 x^{\alpha-1}(1-x)^{\gamma-\alpha-1}\,\mathrm{e}^{zx}\,\mathrm{d}x \ , \tag{19.308}$$

avec $\Re\alpha > 0$ et $\Re(\gamma - \alpha) > 0$.

8. Combien vaut $\Phi(\alpha,\,\gamma,\,z)$ en $z = 0$?

9. En déduire que $\Phi(\alpha,\,\gamma,\,z)$ admet la représentation intégrale[96] :

$$\Phi(\alpha,\,\gamma,\,z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)}\int_0^1 x^{\alpha-1}(1-x)^{\gamma-\alpha-1}\,\mathrm{e}^{zx}\,\mathrm{d}x \ , \tag{19.311}$$

où $\Gamma(z)$ est la fonction d'Euler de deuxième espèce , et avec toujours $0 < \Re\alpha < \Re\gamma$.

10. En posant $L(z) = z^{1-\gamma}M(z)$, montrer que la fonction $M(z)$ satisfait l'équation de Laplace où $\gamma \to 2-\gamma$ et $\alpha \to \alpha - \gamma + 1$. En déduire la forme de l'autre solution linéairement indépendante de l'équation (19.137) (c'est la solution non analytique en $z = 0$, écartée pour des raisons physiques).

19.3.14 À propos du vecteur de Lenz - Runge

1. Montrer que le vecteur de Lenz - Runge défini en (19.235) a des crochets de Poisson nuls avec le Hamiltonien Coulombien.

2. Démontrer les relations de commutation (19.239) - (19.241) de l'opérateur $\vec{R}$ posé en (19.238).

3. Établir l'égalité (19.236).

---

[96] On donne l'intégrale :

$$B(p,\,q) \stackrel{\text{déf}}{=} \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \qquad (\Re p > 0,\ \Re q > 0)\ . \tag{19.309}$$

La fonction $B(p,\,q)$ ainsi définie s'appelle *fonction d'Euler de première espèce* et, tout comme $\Gamma(z)$, se prolonge dans $\mathbb{C}\backslash-\mathbb{N}$ (elle a évidemment les mêmes pôles – simples – que $\Gamma(z)$).

Par ailleurs, on peut montrer que la représentation (19.311) reste vraie même quand γ n'est pas un entier pair, mais n'a toujours évidemment de sens que si :

$$\Re\alpha > 0 \ , \qquad \Re(\gamma-\alpha) > 0 \ . \tag{19.310}$$

19.3.15 Écart à l'interaction de Coulomb : écrantage en loi-puissance. Suppression de la dégénérescence accidentelle

Pour un atome alcalin (un électron sur la couche externe), l'attraction par le noyau de l'électron périphérique est quelque peu écrantée par les électrons internes. Pour rendre compte de cet effet, on corrige le potentiel Coulombien $-\frac{Ze'^2}{r}$ par un terme déstabilisant en r^{-2} ; pour la commodité, on écrit ainsi l'énergie potentielle sous la forme :

$$V(r) = -\frac{Ze'^2}{r} + C\frac{\hbar^2}{2\mu r^2} \ , \qquad (19.312)$$

où C est une constante positive sans dimension.

1. En reprenant pas à pas l'analyse de la sous-section 19.1.2, montrer que la fonction radiale $R(r)$ se comporte comme $r^{\beta-1}$ quand $r \to 0$, et trouver l'exposant β en fonction de l et C.

2. De même, en suivant l'argumentation du paragraphe p. 803, mais sans faire de calcul, montrer que l'énergie est maintenant donnée par :

$$E_{Nl} = -\frac{\mu Z^2 e'^4}{2\left[N + \frac{1}{2}\left(1 + \sqrt{(2l+1)^2 + C}\right)\right]^2 \hbar^2} \ , \qquad (19.313)$$

où $N \in \mathbb{N}$. Commenter.

physique

Chapitre 20

Le spin

Le premier but de ce chapitre est d'introduire le spin d'une particule sur des bases physiques et d'examiner les conséquences de son existence sur le magnétisme atomique, fournissant au passage une illustration de la théorie du moment cinétique développée antérieurement. Les équations de Klein - Gordon et de Dirac sont ensuite établies sur des arguments élémentaires. L'analyse critique de la théorie de Dirac permet de mesurer à la fois ses immenses succès et son rôle historique pour l'émergence de la Théorie quantique des champs. Finalement, la limite faiblement relativiste est examinée, permettant la validation théorique du Hamiltonien de Pauli.

20.1 Insuffisances de la description par une seule fonction d'onde

Jusqu'à présent, l'électron a été considéré comme un objet ponctuel et sans aucune structure ou degré de liberté interne[1]. Dans cette optique, en représentation-q pour fixer les idées, tous les degrés de liberté quantiques sont exprimables à l'aide des seules coordonnées d'espace (x, y et z par exemple). Toute l'information sur l'état du système à l'instant t est alors réputée entièrement contenue dans la connaissance de la fonction d'onde $\Psi(x,\, y,\, z,\, t)$, satisfaisant l'équation de Schrödinger $i\hbar\frac{\partial}{\partial t}\Psi = H\Psi$, un état initial étant supposé donné pour avoir une réponse unique au problème posé.

[1] De façon assez énigmatique, dès que l'on sort du cadre de la relativité de Galilée en introduisant une valeur *finie* c pour la vitesse de la lumière, il est possible pour toute particule de masse m de définir une échelle de longueur *finie*, sa longueur d'onde Compton $\frac{\hbar}{mc}$. Ceci est vrai... même pour un objet réputé *ponctuel* comme l'électron.

Une telle description est insuffisante, comme on va le voir. Cette *non-complétude* provient des preuves expérimentales démontrant qu'une particule (réputée élémentaire) telle que l'électron possède un moment magnétique *propre*, indépendamment de tout mouvement de rotation dans l'espace autour d'un centre. L'existence de ce moment magnétique entraîne à son tour celle, inévitable, d'un moment cinétique propre, ou intrinsèque, qui a été baptisé *spin* par les physiciens car on croyait au début que ce degré de liberté était lié à une rotation de la particule sur elle-même[2]. Ce degré de liberté est "interne" – bien que l'électron continue à être considéré comme une particule ponctuelle et élémentaire ; c'est, au même titre que la charge ou la masse, un attribut intrinsèque, donné une fois pour toutes. Il s'avère impossible de donner du spin une image classique ; se représenter l'électron comme une petite bille de rayon non-nul qui tourne sur elle-même conduit à des absurdités[3]. On verra par la suite que le moment cinétique total s'exprime comme la somme des moments cinétiques orbital et de spin ; c'est pourquoi on peut dire que le spin d'une particule est son moment cinétique dans le référentiel où elle est au repos ([56], p. II-4) – d'où les particularités du spin du photon ([57], § 6), qui n'est au repos dans aucun référentiel.

L'hypothèse du spin de l'électron a été formulée par Uhlenbeck et Goudsmit en 1925 pour rendre compte pragmatiquement des spectres des atomes complexes[4] ; alliée au Principe de Pauli, elle est le fondement de la classification périodique des éléments de Mendeleiev. Bien au-delà de ces vertus explicatives, cette hypothèse a des implications et des conséquences considérables : le spin permet de distinguer deux classes de particules (les fermions et les bosons) ; c'est cette distinction qui permet, finalement, de comprendre ni plus ni moins que la stabilité de la matière [58] et, notamment, d'assurer l'existence de la limite thermodynamique quand des particules de charges contraires sont en présence.

Les preuves de l'existence d'un moment magnétique propre de l'électron sont multiples. L'expérience la plus importante – sur ce point comme pour d'autres – est certainement celle de Stern et Gerlach (voir Tome I, chapitre 11). Dans sa version historique, elle impliquait des atomes d'argent dans l'état fondamental dont la symétrie sphérique exclut l'existence d'un moment cinétique de type orbital non-nul en moyenne (voir chapitre 18) et, par voie de conséquence, celle d'un moment magnétique permanent. Un moment cinétique orbital *fini*, à lui seul, aurait pu expliquer l'existence d'un moment magnétique engendré par une boucle de courant, mais, de toute façon, le nombre de taches observées sur l'écran est incompatible avec un moment cinétique orbital : pour l'atome d'Ag, on observe deux taches (nombre *pair*) alors que, comme on l'a vu au chapitre 18, un moment cinétique orbital conduit à un nombre *impair* de valeurs possibles pour la projection le long d'un axe donné. L'apparition d'un nombre pair de taches exclut définitivement un moment magnétique d'origine orbital, lié classiquement au moment cinétique par le facteur gyromagnétique ordinaire.

Une autre preuve est fournie par l'effet Zeeman anormal. L'effet Zeeman[5] désigne

[2]On parle du spin de la Terre pour évoquer sa révolution diurne. Le mot *spineur* a été utilisé par Élie Cartan en 1913 pour désigner des objets ayant certain comportement par rotation – voir p. 856.

[3]Par exemple, on trouve qu'un point situé à la périphérie de l'électron a une vitesse linéaire très supérieure à c (voir Tome I, problème 11.3.2)

[4]À propos de l'introduction première du spin par Uhlenbeck et Goudsmit (1925), voir Tome I, p. 379.

[5]Pieter Zeeman (1865-1943) partagea le Nobel 1902 avec Lorentz pour la découverte de cet effet. Très

généralement l'éclatement des raies spectrales d'un atome quand on le soumet à un champ magnétique ; en présence du champ, chaque raie se décompose en plusieurs raies très voisines (*splitting*), dont l'écart en fréquence par rapport à la raie unique observée en l'absence de champ est, en champ faible, simplement proportionnel au module du champ appliqué. Les premières observations faisaient état d'une raie en champ nul se décomposant en trois raies très proches, plus généralement en un nombre impair de raies, dont une explication a même pu être fournie dans le cadre de l'Ancienne Théorie des Quanta (voir Tome I, p. 222). En revanche, des observations ultérieures ont révélé la possibilité d'un éclatement en un nombre *pair* de raies, phénomène à nouveau incompréhensible tant que seuls des moments cinétiques orbitaux sont en jeu.

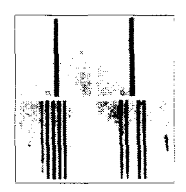


Figure 20.1: Photo prise par Zeeman lui-même de l'effet portant son nom.

Enfin, la structure fine des raies spectrales est une autre conséquence du moment magnétique propre de l'électron, et sert de point d'ancrage à un argument théorique définitif montrant que seule l'existence du spin permet de sortir d'une grave impasse. On sait qu'un moment magnétique $\vec{\mu}$ qui se déplace à la vitesse $\vec{v}$ dans un champ électrique $\vec{\mathcal{E}}$ se couple, par son seul mouvement[6], à ce champ électrique ; l'interaction est donnée par $\mu\vec{v}\times\vec{\mathcal{E}}$. Si le champ est à symétrie sphérique, il est forcément de la forme $\vec{\mathcal{E}} = f(r)\vec{r}$; on trouve alors la combinaison $\vec{r}\times\vec{v}$, proportionnelle au moment cinétique orbital $\vec{L}$, d'où finalement dans le Hamiltonien un terme du genre :

$$\boxed{V_{\text{magn}} \stackrel{\text{déf}}{=} a(r)\vec{L}.\vec{\mu}} \tag{20.1}$$

Ce résultat majeur peut s'obtenir comme suit à partir d'arguments purement classiques. Dans le repère de l'électron, celui-ci voit le noyau (une particule chargée !) tourner autour de lui à la vitesse $\vec{v}_{\text{N}}$, produisant le champ magnétique $\vec{\mathcal{B}}$ ($\rho(\vec{r}\,') = \delta(\vec{r}\,')$ puisque le noyau est pris ponctuel et situé à l'origine) :

$$\vec{\mathcal{B}}(\vec{r}) = \frac{\mu_0}{4\pi}\int_{\mathbb{R}^3} \mathrm{d}^3r'\,\frac{\rho(\vec{r}\,')\vec{v}_{\text{N}}\times(\vec{r}-\vec{r}\,')}{\|\,(\vec{r}-\vec{r}\,')\,\|^3} = \frac{\mu_0}{4\pi}\,\frac{Z|e|\,\vec{v}_{\text{N}}\times\vec{r}}{r^3}\ . \tag{20.2}$$

tôt intéressé par la physique, il fit néanmoins des études d'humanités avant de préparer sa thèse sous la direction de Kamerlingh Onnes et Lorentz. Il est aussi l'un des concepteurs des *Travaux du Zuidersee*, entreprise pharaonique destinée à étendre les terres agricoles (*polders*), et à protéger les Pays-Bas des inondations.

[6]Tout comme une charge électrique q en mouvement dans un champ magnétique $\vec{\mathcal{B}}$ ressent une force $\vec{F} = q\vec{v}\times\vec{\mathcal{B}}$.

Dans le repère du noyau, la vitesse $\vec{v}$ de l'électron est égale à $-\vec{v}_{\rm N}$; le champ magnétique est donc aussi :

$$\vec{\mathcal{B}}(\vec{r}\,) = -\frac{\mu_0}{4\pi}\,\frac{Z|e|\,\vec{v}\times\vec{r}}{r^3} \equiv \frac{\mu_0}{4\pi}\,\frac{1}{m}\,\frac{Z|e|}{r^3}\,\vec{L}\ , \tag{20.3}$$

où $\vec{L} = \vec{r}\times m\vec{v}$ est le moment cinétique orbital de l'électron en rotation autour du noyau. En désignant par $U(r) = \frac{Z|e|}{4\pi\varepsilon_0 r}$ le potentiel électrostatique du noyau, (20.3) s'écrit[7] :

$$\vec{\mathcal{B}}(\vec{r}\,) = -\frac{1}{mc^2}\,\frac{1}{r}\,\frac{\mathrm{d}U}{\mathrm{d}r}\,\vec{L}\ . \tag{20.4}$$

L'électron se couple à ce champ magnétique par l'intermédiaire de son moment magnétique supposé ; si l'on poursuit le raisonnement dans le cadre classique, le terme d'interaction est :

$$-\vec{\mu}.\vec{\mathcal{B}} = +\frac{1}{mc^2}\,\frac{1}{r}\,\frac{\mathrm{d}U}{\mathrm{d}r}\,\vec{L}.\vec{\mu}\ , \tag{20.5}$$

qui est bien de la forme anticipée (20.1). En réalité, en raison d'un effet subtil appelé *précession de Thomas*[8], dû à la "contraction du temps", l'expression correcte de $a(r)$ contient un facteur 1/2 additionnel[9] :

$$\boxed{a(r) = +\frac{1}{2mc^2}\,\frac{1}{r}\,\frac{\mathrm{d}U}{\mathrm{d}r}} \tag{20.6}$$

L'origine physique de $V_{\rm magn}$ justifie que l'on donne à ce terme le nom d'*interaction spin-orbite*, puisqu'il est une conséquence directe de l'*orbite* décrite par l'électron nanti de son moment magnétique. Ce couplage produit des (petites) variations de l'énergie des différents états ; en effet, en remplaçant r par a_0 dans (20.5) et en admettant que $\vec{\mu}$ est lié à un moment cinétique par un facteur gyromagnétique du genre $\frac{|e|}{m}$ (à des facteurs près)[10], on trouve, en ordre de grandeur :

$$V_{\rm magn} \sim \frac{1}{mc^2}\,\frac{1}{a_0^3}\,\frac{Z|e|}{4\pi\varepsilon_0}\hbar\frac{|e|}{2m}\hbar \sim \left(\frac{\hbar}{mca_0}\right)^2 E_{n=1} \tag{20.7}$$

où $E_{n=1}$ est l'énergie fondamentale de l'atome d'hydrogène $\sim \frac{e'^2}{a_0}$. Compte tenu de la définition de la première orbite de Bohr, $a_0 = \frac{\hbar^2}{me'^2}$, il vient :

$$\boxed{V_{\rm magn} \sim \alpha^2 E_{n=1} \simeq 7\times 10^{-4}\,\mathrm{eV}}\ , \tag{20.8}$$

α étant toujours la constante de structure fine, définie en (19.173). Les déplacements en énergie varient d'un état à l'autre, mais ils restent petits vis-à-vis des énergies purement

[7] $\varepsilon_0\mu_0 c^2 = 1$.

[8] Llewellyn Hilleth Thomas (1903-1992) était physicien et mathématicien. Il est connu pour avoir mis en évidence la précession particulière portant son nom [59], mais également pour avoir introduit une description statistique de l'atome, appelée depuis *modèle de Thomas - Fermi*, repris et développé par Dirac et Weizsäcker. Ce modèle a inspiré les méthodes plus récentes de *fonctionnelle de densité*.

[9] voir sous-section 20.7, et notamment l'expression (20.245).

[10] L'argument est visiblement d'inspiration classique ; il est cependant recevable sur un strict plan d'homogénéité et en ne jouant qu'avec les constantes fondamentales disponibles.

électrostatiques (voir (20.16)). En tout cas, combinés avec la règle de Bohr $E_\text{f} - E_\text{i} = h\nu$, ils donnent lieu à l'apparition d'une structure dans les raies spectrales : avec une résolution médiocre, on ne voit qu'une seule raie ; en regardant mieux, on constate que chaque raie est le plus souvent constituée de plusieurs raies très voisines : c'est ce que l'on appelle la structure fine[11]. Sur des considérations *apparemment* sans rapport, Sommerfeld fut le premier a en fournir une explication remarquable pour le cas particulier de la raie d'émission H_α ($n = 3 \rightarrow n = 2$) de la série de Balmer, effectuant un calcul relativiste dans le cadre de l'Ancienne Théorie des Quanta (voir Tome I, sous-section 7.3.3). En se ralliant à l'explication théorique historiquement légitime (voir section 20.5.1) de l'émergence du spin de l'électron, on peut dire que ceci n'est pas tout à fait un hasard.

L'existence du moment magnétique propre de l'électron étant admise sous la pression des faits expérimentaux, un argument[12] permet alors de comprendre que, *nécessairement*, l'électron possède alors un moment cinétique *propre*. En effet, la présence d'un terme comme V_magn dans le Hamiltonien H entraîne que le vecteur $\vec{L}$ n'est plus une constante du mouvement : comme les différentes composantes de $\vec{L}$ ne commutent pas entre elles, le commutateur $[L_u,\, \vec{L}.\vec{\mu}]$ n'est pas nul. Ainsi, en l'état actuel des choses, le moment cinétique d'un système isolé comme l'atome n'est pas une constante du mouvement, en violation flagrante de l'invariance liée à l'isotropie de l'espace.

La difficulté se résout d'elle-même si, par analogie avec le cas orbital, on fait l'hypothèse que le moment magnétique intrinsèque avéré est inévitablement associé à un moment cinétique propre, noté $\vec{S}$, dont l'existence est postulée, et tel que $\vec{\mu} \propto \vec{S}$, plus précisément $\vec{\mu} = \gamma_\text{e}\vec{S}$ où γ_e est un certain facteur gyromagnétique. En raison du caractère *intrinsèque* du moment cinétique de spin, il n'y a pas lieu de considérer ce facteur gyromagnétique comme universel : il est *a priori* propre à l'électron d'où l'indice $_\text{e}$; dans ces conditions, l'interaction spin-orbite V_magn (20.1) prend la forme :

$$V_\text{magn} = A(r)\,\vec{L}.\vec{S}\ , \qquad A(r) = \gamma_\text{e}\frac{1}{2mc^2}\,\frac{1}{r}\,\frac{\text{d}U}{\text{d}r}\ . \tag{20.9}$$

Une fois ceci admis, il est alors naturel de *définir* le moment cinétique *total*, noté $\vec{J}$, par la combinaison :

$$\boxed{\vec{J} \overset{\text{déf}}{=} \vec{L} + \vec{S}} \tag{20.10}$$

qui, en tant que somme de deux moments cinétiques, est un moment cinétique (voir sous-section 18.5.1). Si $\vec{J}$ est le *bon* moment cinétique total, il doit être une constante du mouvement ; montrons qu'il en est bien ainsi. On a :

$$\vec{L}.\vec{S} = \frac{1}{2}\left[(\vec{L}+\vec{S})^2 - \vec{L}^{\,2} - \vec{S}^{\,2}\right] = \frac{1}{2}(\vec{J}^2 - \vec{L}^{\,2} - \vec{S}^{\,2}) \tag{20.11}$$

d'où :

$$[\vec{J},\, \vec{L}.\vec{S}] = \frac{1}{2}[\vec{J},\, \vec{J}^2 - \vec{L}^{\,2} - \vec{S}^{\,2}] = -\frac{1}{2}[\vec{L}+\vec{S},\, \vec{L}^{\,2} + \vec{S}^{\,2}]\ . \tag{20.12}$$

[11]La structure hyperfine (impliquant des effets encore plus petits) résulte d'un couplage du même type avec le spin nucléaire (voir section 27.4). Cette fois c'est le noyau qui, par son spin (donc son moment magnétique), se couple au champ magnétique créé par la rotation des électrons ; à ce couplage, il convient d'ailleurs en toute rigueur d'ajouter l'interaction dipôle - dipôle entre les moments magnétiques.

[12]voir [4] p. 255.

À ce point précis apparaît la nécessité de connaître les règles de commutation de $\vec{L}$ et $\vec{S}$. $\vec{L}$ s'exprime avec des grandeurs dynamiques plus "fondamentales" selon $\vec{L} = \vec{r} \times \vec{p}$; au contraire, $\vec{S}$ est une caractéristique propre supplémentaire introduite par nécessité. Dans ces conditions, il est logique d'admettre que, se référant à des degrés de liberté différents, ces deux opérateurs commutent :

$$\boxed{[L_u,\, S_v] = 0 \qquad \forall u,\, v = x,\, y,\, z} \tag{20.13}$$

Cela étant, (20.12) donne $[\vec{J},\, \vec{L}.\vec{S}] = -\frac{1}{2}[\vec{L},\, \vec{L}^{\,2}] - \frac{1}{2}[\vec{S},\, \vec{S}^{\,2}] = 0$: $\vec{J}$ est bien le moment cinétique qui, même en présence d'interaction spin-orbite, est une constante du mouvement. Ainsi, il est impossible d'admettre l'existence d'un moment magnétique propre de l'électron non lié à un moment cinétique, tout en voulant préserver la conservation *obligatoire* du moment cinétique d'un système isolé. Au total, l'existence avérée expérimentalement du moment *magnétique* de l'électron, alliée à la symétrie imposée par l'isotropie, force à admettre l'existence du moment *cinétique* propre de l'électron.

Étant acquis que c'est bien $\vec{J}$ le moment cinétique de l'électron en orbite nanti de son spin, $\vec{L}$ et $\vec{S}$ deviennent des opérateurs vectoriels : il est en effet facile de vérifier que le commutateur $[\vec{V},\, \vec{J}\,]$, avec $\vec{V} = \vec{L}$ ou $\vec{S}$, est égal à ce qui a été énoncé dans la théorie du moment cinétique pour élire un opérateur à trois composantes en tant qu'opérateur vectoriel[13] (voir éq. (17.117)).

La relation classique entre moment magnétique et moment cinétique a été rappelée dans le Tome I (voir éq. (11.8)) ; pour l'électron, elle s'écrit :

$$\vec{\mu} = \frac{e}{2m}\,\vec{L} \quad ; \tag{20.14}$$

la charge étant négative, les deux vecteurs sont dirigés en sens contraires. Plus généralement, pour une particule de masse M et de charge q, la relation classique est :

$$\boxed{\vec{\mu} = \frac{q}{2M}\,\vec{L} \equiv \gamma\,\vec{L}} \tag{20.15}$$

ce qui définit le facteur gyromagnétique γ. Comme l'unité fondamentale de moment cinétique est $\hbar$, il est d'usage, pour l'électron par exemple, d'écrire le moment magnétique d'origine orbitale sous la forme $\vec{\mu} = \frac{e\hbar}{2m}\,\frac{\vec{L}}{\hbar}$, écriture où apparaît spontanément le *magnéton de Bohr*[14], μ_{B} :

$$\boxed{\mu_{\text{B}} \overset{\text{déf}}{=} \frac{e\hbar}{2m} \simeq -0,927\,4 \times 10^{-23}\,\text{Joule/Tesla} \simeq -5,789 \times 10^{-5}\,\text{eV/Tesla}} \tag{20.16}$$

et alors $\vec{\mu} = \mu_{\text{B}}\hbar^{-1}\vec{L}$. On introduit aussi le magnéton nucléaire, μ_{N}, relatif au proton :

$$\boxed{\mu_{\text{N}} \overset{\text{déf}}{=} \frac{|e|\hbar}{2M_{\text{p}}} \simeq 5,051 \times 10^{-27}\,\text{Joule/Tesla} \simeq 3,153 \times 10^{-8}\,\text{eV/Tesla}} \tag{20.17}$$

[13] Comme $[\vec{L},\, \vec{S}\,] = 0$, aucun des deux opérateurs $\vec{L}$ et $\vec{S}$ n'est un opérateur vectoriel vis-à-vis de l'autre.

[14] On adopte partout la convention suivant laquelle μ_{B} est *négatif* ; il semble que ce soit l'usage le plus répandu ([4], [25], etc.).

Ces valeurs numériques montrent la petitesse des effets magnétiques dans la matière ordinaire, et pour les expériences facilement réalisables au laboratoire ($\mathcal{B}$ de l'ordre de quelques Teslas, au plus).

Les relations précédentes reposent sur des notions classiques et se réfèrent automatiquement à un moment magnétique issu d'un moment cinétique orbital. Étant convaincu de l'existence d'un moment magnétique intrinsèque, on pose naturellement – sur de simples considérations dimensionnelles – des relations analogues entre le moment cinétique de spin et le moment magnétique révélé par l'expérience. Comme le moment magnétique lié au spin ne résulte pas d'une révolution géométrique, il peut apparaître – outre le rapport charge/masse dimensionnellement nécessaire – un facteur numérique que l'on espère d'ordre unité sinon le calibrage serait malhabile ; pour le moment magnétique de spin, on pose ainsi pour l'électron :

$$\boxed{\vec{\mu}_{\rm S} = g_{\rm e}\,\frac{e}{2m}\,\vec{S} = g_{\rm e}\,\mu_{\rm B}\,\hbar^{-1}\,\vec{S} \equiv \gamma_{\rm e}\,\vec{S}} \qquad (20.18)$$

Dans (20.18), $g_{\rm e}$ est un nombre pur, dont l'expérience montre qu'il est très voisin[15] de 2 ; tout comme $\mu_{\rm B}$, $\gamma_{\rm e}$, facteur gyromagnétique propre de l'électron, est négatif. Avec ces définitions, l'interaction spin-orbite (20.9) prend la forme définitive :

$$\boxed{V_{\rm magn} = A(r)\,\vec{L}.\vec{S} \qquad A(r) = g_{\rm e}\frac{1}{4m^2c^2}\,\frac{1}{r}\,\frac{{\rm d}V}{{\rm d}r} \equiv \frac{g_{\rm e}e'^2}{4m^2c^2}\,\frac{1}{r^3}} \qquad (20.20)$$

Pour le proton[16], de façon analogue :

$$\boxed{\vec{\mu}_{\rm S} = g_{\rm p}\,\frac{|e|}{2M_{\rm p}}\,\vec{S} = g_{\rm p}\,\mu_{\rm N}\,\hbar^{-1}\,\vec{S} \equiv \gamma_{\rm p}\,\vec{S}} \qquad (20.21)$$

avec $g_{\rm p} \simeq 5,585\,694\,713$ [61]. Pour le neutron[17], $g_{\rm n} \simeq -3,826\,085\,45$, [62] conférant au neutron un moment magnétique égal à 1,913 15 magnéton nucléaire (tout comme l'électron, le neutron a ses deux moments de spin et magnétique opposés). L'existence d'un moment magnétique pour le neutron (globalement *neutre*[18]) peut sembler mystérieuse au vu de la relation classique (20.15) – qu'il vaut mieux dans ce cas lire à l'envers,

[15] La théorie de Dirac donne strictement $g_{\rm e} = 2$ (voir section 20.5.1), expérimentalement, $g_{\rm e} \gtrsim 2$. L'écart fini $g_{\rm e} - 2$ résulte de corrections électrodynamiques ; son calcul est l'un des enjeux de l'Électrodynamique quantique, dont le succès à cet égard est proprement fabuleux ; les valeurs théoriques et expérimentales coïncident sur une douzaine de chiffres significatifs :

$$\boxed{g_{\rm th} = 2,002\,319\,304\,362\,2\ \pm 6\times 10^{-11} \qquad g_{\rm exp} = 2,002\,319\,304\,376 \pm 6\times 10^{-11}} \qquad (20.19)$$

[16] également de spin $S = 1/2$, et de masse 938,258 MeV. Selon l'habitude, dire qu'une particule est de spin S c'est donner la valeur du nombre quantique donnant le carré du module de son moment cinétique intrinsèque ; pour "une particule de spin S", $\|\vec{S}\|^2 = S(S+1)\,\hbar^2$.

[17] Le neutron a aussi un spin $S = 1/2$; sa masse est très légèrement supérieure à celle du proton : $M_{\rm n} \simeq 939,553$ MeV. Le neutron (libre) est instable (durée de vie de l'ordre de la dizaine de minutes), alors que le proton est stable, pour autant qu'on sache.

[18] La charge du neutron est expérimentalement inférieure à 2×10^{-22} charge élémentaire [62].

partant de l'idée que c'est le spin qui est l'attribut intrinsèque fondamental (tout comme une charge, une masse).

C'est le spin, quand il est demi-entier, qui permet d'expliquer le nombre pair de raies Zeeman anormales et le nombre pair d'impacts sur l'écran dans l'expérience de Stern et Gerlach, étant entendu que l'on observe aussi des cas où ces nombres sont impairs, bien qu'il s'agisse toujours d'électrons au sein d'atomes. De fait, pour un atome complexe (et en ne considérant que les électrons) la parité du nombre de signaux résulte de l'addition de tous les moments cinétiques, orbitaux et de spin, de tous les électrons. Même si $\vec{L}$ est non-pertinent pour le cas examiné (parce que l'atome est, du point de vue orbital, dans un état sphérique $\vec{L} = 0$), le spin à lui seul peut donner soit des nombres pairs, soit des nombres impairs : tout dépend de la *parité* du nombre d'électrons dans l'atome (voir chapitre 27). La théorie générale du moment cinétique permet d'affirmer que les atomes à nombre pair (resp. impair) d'électrons ont un spin entier (resp. demi-entier)[19].

Le spin S d'une particule est toujours un nombre demi-entier *ou* entier, c'est un fait d'expérience. Le caractère entier ou demi-entier du spin définit deux grandes classes de particules, les *bosons*[20] (spin entier) et les *fermions*[21] (spin demi-entier), obéissant à des statistiques très différentes, d'où l'existence d'une relation[22] appelée *théorème spin-statistique*, énoncée par Pauli en 1940 [65]. Par ailleurs, si S peut prendre des valeurs relativement élevées (on connaît des noyaux qui, dans leur état fondamental, ont un spin égal à 11/2), il semble toutefois difficile de concevoir des spins arbitrairement grands à l'instar des moments cinétiques orbitaux[23] ; le spin ne se prête pas à la limite quasi-classique, ce qui n'est finalement pas très surprenant : de nature essentiellement quantique, il s'évanouit dans cette limite, tout comme les particularités liées aux particules identiques (voir chapitre 22) ; une limite telle que $\hbar \to 0$, $S \to +\infty$ avec $\hbar S = \mathrm{C^{ste}}$ n'a pas de sens. Comme mentionné ailleurs (voir Remarque 1 p. 866), si la Physique classique

[19] Cette affirmation laisse de côté le spin nucléaire. Pour l'hydrogène ordinaire ("léger"), le noyau a un spin 1/2 (un proton) et, au total, le spin de l'atome vaut soit 0, soit 1 ; ceci permet, si l'on veut, de distinguer deux "espèces" d'hydrogène. Cette classification n'est utile que pour les problèmes où la structure hyperfine joue un rôle important, ou lorsque l'aspect boson vs. fermion est pertinent.

[20] Ce nom a été choisi en hommage à Satyendranath Bose (~ 1894-1974). Alors qu'il enseignait à l'université de Dacca, Bose envoya à Einstein un très court article... que venait de lui rejeter le *Philosophical Magazine*. C'est Einstein qui en a fit la traduction en allemand et le recommanda pour publication dans le *Zeitschrift für Physik* [63] : le *condensat de Bose - Einstein* était révélé à la communauté des physiciens ! Bose fut le premier à reconnaître que la statistique de Maxwell - Boltzmann ne pouvait pas être appliquée dans le cadre quantique.

[21] À nouveau, l'appellation est un hommage, cette fois à Fermi. Enrico Fermi (1901-1954), Prix Nobel en 1938, a dû s'exiler en raison des lois antijuives promulguées par le régime de Mussolini. Après avoir travaillé à Chicago sur la première pile nucléaire, il fut l'un des artisans du Projet Manhattan. Fermi était réputé pour se poser les problèmes en termes mathématiques très simples, avant d'en fournir une solution très élaborée une fois sa réflexion achevée. Il est mort prématurément d'un cancer de l'estomac, certainement lié à son exposition sans précaution durant ses expériences de Chicago.

Une très belle biographie de Fermi est due à Emilio G. Segrè [64], autre Prix Nobel (1959, avec Owen Chamberlain pour leur mise en évidence de l'anti-proton), autre artisan du *Manhattan Project*.

[22] À propos du statut de cette relation (règle empirique, théorème, ...), voir [25], XIV.C.1. – pour une revue assez récente, voir [66].

[23] Dans la limite des très grands nombres quantiques pour le moment orbital, la fonction de Brillouin rencontrée dans l'étude du paramagnétisme se confond pratiquement avec la fonction classique de Langevin (voir Tome I, problème 11.3.1).

ignore le spin, c'est parce que, au fond, elle peut se passer des nombres complexes qui ne sont jamais, dans le cadre classique qu'une *commodité technique* dont on peut toujours se dispenser[24].

Revenons au cas de l'électron. Les deux valeurs possibles révélées par une mesure de S_z sont donc $\pm\hbar/2$, associées aux deux valeurs possibles d'un nombre quantique binaire $M_S = \pm 1/2$, lui-même associé à un nombre $S = 1/2$ donnant le module d'un moment cinétique ($\|\vec{S}\| = \sqrt{S(S+1)}\,\hbar = \sqrt{3}\,\hbar/2$). Une description complète de l'état de l'électron contient nécessairement une fonction d'onde donnant comme d'habitude la densité de probabilité de présence, mais prenant également en compte le degré de liberté de spin, d'où la notation $\Psi(\vec{r},\, M_S,\, t)$. Si les coordonnées d'espace peuvent prendre des valeurs continues, en revanche la variable de spin est essentiellement discrète et toute mesure de S_z ne peut, donner ici que l'une des deux valeurs ci-dessus. En maintenant l'interprétation usuelle, la quantité $|\Psi(\vec{r},\, M_S,\, t)|^2\,\mathrm{d}^3r$ est ainsi la probabilité de présence autour du point fixé par le vecteur $\vec{r}$ et avec la valeur $M_S\hbar$ pour le spin. La condition de normalisation des probabilités introduit comme toujours une sommation, qui porte non seulement sur les degrés orbitaux (sommation continue, c'est-à-dire intégration) et mais également sur les degrés de spin (sommation discrète) :

$$\int_{\mathbb{R}^3} |\Psi(\vec{r},\, +\frac{1}{2},\, t)|^2\,\mathrm{d}^3r + \int_{\mathbb{R}^3} |\Psi(\vec{r},\, -\frac{1}{2},\, t)|^2\,\mathrm{d}^3r = 1 \tag{20.22}$$

exprimant notamment le fait que l'on épuise toutes les possibilités du spin en sommant sur les deux seules valeurs possibles. L'équation (20.22) implique que si l'on fait une mesure simultanée de position *et* de spin – plus précisément de S_z –, on peut trouver l'électron près de $\vec{r}$ avec le spin $+\hbar/2$ *ou* avec le spin $-\hbar/2$. En termes plus formels, (20.22) se récrit :

$$\sum_{M_S=\pm\frac{1}{2}} \int_{\mathbb{R}^3} |\Psi(\vec{r},\, M_S,\, t)|^2\,\mathrm{d}^3r = 1\ . \tag{20.23}$$

En tout état de cause, l'électron n'a plus une mais deux fonctions d'onde, une pour chaque valeur[25] de S_z. De la même façon, pour une particule de spin $S = 1$, M_S prend les trois valeurs $0, \pm 1$ et on écrit :

$$\sum_{M_S=0,\,\pm 1} \int_{\mathbb{R}^3} \Psi(\vec{r},\, M_S,\, t)|^2\,\mathrm{d}^3r = 1\ . \tag{20.24}$$

Il y a maintenant trois fonctions d'onde, et ainsi de suite : une particule de spin S aura $2S+1$ fonctions d'onde ou, mieux, une fonction d'onde à $2S+1$ composantes.

La notation précédente n'est pas forcément la meilleure, car s'agissant d'une variable prenant des valeurs discrètes, il est tout aussi légitime (et même plus naturel) de mettre M_S en indice et d'écrire $\Psi_{M_S}(\vec{r},\, t)$ au lieu de $\Psi(\vec{r},\, M_S,\, t)$. Enfin, il est commode d'utiliser une notation matricielle, rangeant en colonne les différentes fonctions correspondant aux valeurs possibles de la variable discrète M_S. Ainsi, pour l'électron,

[24] À ce sujet, relire les articles de Schrödinger [67], notamment pp. 66 et 96.

[25] ou deux combinaisons linéaires indépendantes des deux états propres de S_z.

on admettra désormais que toute l'information au sens de la Mécanique quantique est contenue dans une matrice colonne à deux lignes contenant les deux composantes et appelée *spineur*[26] – terme proposé par Élie Cartan[27] peu après 1910, plus d'une décennie avant la découverte du spin des particules –, et noté :

$$\begin{bmatrix} \Psi_{+\frac{1}{2}}(\vec{r},\, t) \\ \Psi_{-\frac{1}{2}}(\vec{r},\, t) \end{bmatrix} , \tag{20.25}$$

ou, encore plus simplement :

$$\begin{bmatrix} \Psi_{+}(\vec{r},\, t) \\ \Psi_{-}(\vec{r},\, t) \end{bmatrix} ; \tag{20.26}$$

l'état d'une particule de spin $S = 1/2$ est ainsi maintenant décrit par un certain *doublet*. Pour une particule de spin $S = 1$, on posera de même :

$$\begin{bmatrix} \Psi_{+1}(\vec{r},\, t) \\ \Psi_{0}(\vec{r},\, t) \\ \Psi_{-1}(\vec{r},\, t) \end{bmatrix} , \tag{20.27}$$

ou, quand aucune ambiguïté n'est à craindre :

$$\begin{bmatrix} \Psi_{+}(\vec{r},\, t) \\ \Psi_{0}(\vec{r},\, t) \\ \Psi_{-}(\vec{r},\, t) \end{bmatrix} . \tag{20.28}$$

La fonction d'onde a maintenant trois composantes, formant un *triplet* se comportant comme un *vecteur* par rotation ; pour un spin S, il y a $2S + 1$ composantes, formant un $2S + 1$-plet. Seules les particules de spin nul restent décrites complètement par une seule fonction d'onde.

La notation matricielle ci-dessus se transcrit immédiatement en notation de Dirac ; pour une particule de spin $S = 1/2$, le vecteur d'état complet (i.e. incluant la variable de spin) s'écrit :

$$\boxed{|\Psi(\vec{r},\, t)\rangle = \Psi_{+\frac{1}{2}}(\vec{r},\, t)| + \frac{1}{2}\rangle + \Psi_{-\frac{1}{2}}(\vec{r},\, t)| - \frac{1}{2}\rangle} \tag{20.29}$$

$| \pm \frac{1}{2}\rangle$ désignent les états propres de S_z (et de $\vec{S}^{\,2}$) et forment une base de l'espace des états de spin ; les $\Psi_{\pm\frac{1}{2}}(\vec{r},\, t)$ sont les coefficients du développement de l'état global. On peut d'ailleurs monter d'un cran dans l'usage de la notation de Dirac en identifiant :

$$\langle\vec{r},\, M_S|\Psi(t)\rangle \equiv \Psi_{M_S}(\vec{r},\, t) \ . \tag{20.30}$$

[26]C'est ainsi que l'on appelle un objet se comportant d'une façon bien déterminée par rotation – voir plus loin section 20.3. Ces spineurs sont les objets sur lesquels agissent les matrices 2×2 constituant le groupe SU(2) (*special unitary group*), qui est une représentation *à une phase près* du groupe des rotations [12].

[27]Élie Cartan (1869-1951), est l'un des plus grands mathématiciens français de son époque. Il est surtout connu pour son travail sur les groupes et les algèbres de Lie. Le terme proposé par Cartan trouve son origine dans le fait que, précisément, les spineurs se comportent d'une certaine façon quand on effectue une *rotation*.

$|\vec{r}, M_S\rangle$ désigne alors un état produit tensoriel espace ⊗ spin. Cela étant, les relations habituelles se généralisent immédiatement ; par exemple, la fermeture s'écrit :

$$\sum_{M_S=-S}^{+S} \int_{\mathbb{R}^3} |\vec{r}, M_S\rangle\langle\vec{r}, M_S|\, \mathrm{d}^3r = \mathbf{1} \ . \tag{20.31}$$

Si $\{\psi_n(\vec{r})\}$ désigne une base complète pour les variables d'espace, par exemple la base propre d'une observable, on peut noter $\Psi_{n,\, M_S}$ la matrice unicolonne ayant des zéros partout sauf dans la $(S+1-M_S)^{\mathrm{e}}$ ligne :

$$\Psi_{n,\, M_S}(\vec{r}) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \psi_n(\vec{r}) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \begin{matrix} \\ \\ \\ \\ \longleftarrow \text{ligne } S+1-M_S \\ \\ \\ \\ \end{matrix} , \tag{20.32}$$

et son adjoint par la matrice uniligne obtenue par transposition et conjugaison complexe. La relation de fermeture prend alors la forme :

$$\boxed{\sum_n \Psi^*_{n,\, M_S}(\vec{r})\, \Psi_{n,\, M'_S}(\vec{r}\,') = \delta(\vec{r} - \vec{r}\,')\, \delta_{M_S\, M'_S}} \tag{20.33}$$

l'indice n représentant tous les nombres quantiques spécifiant complètement l'état pour une valeur de S_z donnée. Le produit scalaire de deux vecteurs d'état $|\Psi\rangle$ et $|\Phi\rangle$ d'une particule de spin S est :

$$\langle\Psi|\Phi\rangle = \sum_{M_S=-S}^{+S} \int_{\mathbb{R}^3} \Psi^*_{M_S}(\vec{r})\, \Phi_{M_S}(\vec{r})\, \mathrm{d}^3r \ , \tag{20.34}$$

égalité qui peut aussi s'écrire sous forme matricielle ; par exemple, pour $S = 1/2$, (20.34) est identique à :

$$\langle\Psi|\Phi\rangle = \int_{\mathbb{R}^3} \begin{bmatrix} \Psi^*_+(\vec{r}, t) & \Psi^*_-(\vec{r}, t) \end{bmatrix} \begin{bmatrix} \Phi_+(\vec{r}, t) \\ \Phi_-(\vec{r}, t) \end{bmatrix} \mathrm{d}^3r \ . \tag{20.35}$$

Il convient de signaler une autre notation, commode dans la pratique et couramment utilisée pour un spin $S = 1/2$. Soit les deux fonctions définies comme suit :

$$\alpha(M_S) \stackrel{\text{déf}}{=} \begin{cases} 1 & \text{si } M_S = +1/2 \\ 0 & \text{si } M_S = -1/2 \end{cases} , \qquad \beta(M_S) \stackrel{\text{déf}}{=} \begin{cases} 0 & \text{si } M_S = +1/2 \\ 1 & \text{si } M_S = -1/2 \end{cases} . \tag{20.36}$$

Les valeurs de α et de β reflètent les composantes des vecteurs-colonne représentant les deux états $|\pm\rangle$; en fait, en introduisant les deux *kets* $|\alpha\rangle$ et $|\beta\rangle$, on peut écrire l'identification :

$$\alpha(M_S) \equiv \langle M_S|\alpha\rangle \ , \qquad \beta(M_S) \equiv \langle M_S|\beta\rangle \ ; \tag{20.37}$$

l'état complet d'un électron peut alors se noter :

$$\Psi_{M_S}(\vec{r}, t) = \Psi_{+\frac{1}{2}}(\vec{r}, t)\,\alpha(M_S) + \Psi_{-\frac{1}{2}}(\vec{r}, t)\,\beta(M_S) \ . \tag{20.38}$$

En effet, si on choisit $M_S = +1/2$ au premier membre, on obtient bien l'égalité puisque, au second membre, seul le premier terme subsiste. Ainsi écrit, $\Psi(\vec{r}, M_S;\, t)$ apparaît manifestement comme une combinaison linéaire d'états fabriqués en formant le produit tensoriel des états orbitaux ("d'espace") par les états de spin. Notons que les deux fonctions α et β sont orthogonales :

$$\sum_{M_S=\pm\frac{1}{2}} \alpha(M_S)\,\beta(M_S) = 0 \ . \tag{20.39}$$

Parmi toutes les fonctions du type (20.38), il en est où les variables sont séparées. En effet, supposons que les $\Psi_{\pm\frac{1}{2}}$, sont à un facteur près une seule et même fonction :

$$\Psi_{\pm\frac{1}{2}}(\vec{r}, t) = c_{\pm\frac{1}{2}}\,\Psi(\vec{r}, t) \ ; \tag{20.40}$$

alors, d'après (20.38) :

$$\Psi_{M_S}^{(\text{sep})}(\vec{r}, t) = \Psi(\vec{r}, t)\,[c_{+\frac{1}{2}}\,\alpha(M_S) + c_{-\frac{1}{2}}\,\beta(M_S)] \equiv \Psi(\vec{r}, t)\ \otimes\ \chi(M_S) \tag{20.41}$$

et l'état total est bien un simple produit tensoriel espace $\otimes$ spin (pour la simplicité d'écriture, on omet le plus souvent le symbole $\otimes$). D'après (20.35), la norme se calcule suivant :

$$\langle\Psi|\Psi\rangle = \int_{\mathbb{R}^3} |\Psi(\vec{r}, t)|^2\,\big(|c_{+\frac{1}{2}}|^2 + |c_{-\frac{1}{2}}|^2\big)\,\mathrm{d}^3r \ ; \tag{20.42}$$

l'usage consiste naturellement à normaliser chaque facteur à l'unité :

$$\int_{\mathbb{R}^3} |\Psi(\vec{r}, t)|^2\,\mathrm{d}^3r = 1 \ , \qquad |c_{+\frac{1}{2}}|^2 + |c_{-\frac{1}{2}}|^2 = 1 \ . \tag{20.43}$$

Il n'est pas difficile de se convaincre que l'on peut toujours utiliser comme base de l'espace $\mathcal{E}$ des états du genre (20.41) où les degrés de liberté d'espace et de spin sont séparés. Les états *propres* d'un ECOC choisi d'avance ne seront pas toujours de cette forme, en particulier quand il existe un couplage entre spin et espace[28], mais on peut toujours les représenter en combinaison linéaire d'états factorisés. Dans la suite, l'hypothèse d'une base séparée pourra toujours être faite pour la commodité et l'espace des états $\mathcal{E}$ pourra être considéré comme le produit tensoriel $\mathcal{E}^{(\text{orb})} \otimes \mathcal{E}^{(\text{spin})}$. Le contraire d'un état séparé est un état *intriqué*, de la forme :

$$\Psi^{(\text{intr})}(\vec{r}, t) = \sum_{M_S} \Psi_{M_S}(\vec{r}, t)\chi(M_S) \ , \tag{20.44}$$

[28]D'ailleurs, même en l'absence d'interaction spin-orbite, les états acceptables physiquement pour un ensemble de N particules identiques ne sont presque jamais de cette forme. On verra par la suite (chapitre 22) comment, pour un ensemble de fermions, le postulat d'antisymétrisation entraîne que, pour $N > 2$, seul l'état "ferro" – tous les spins alignés – se présente sous la forme séparée, quelle que soit la composante M_S $(-N/2 \leq M_S \leq +N/2)$. Au contraire, pour le cas $N = 2$, tous les états physiques, dans l'approximation électrostatique, sont séparés.

la partie orbitale Ψ_{M_S} n'étant pas la même pour les différentes fonctions de spin. On peut montrer que, dans un repère de Lorentz donné, un état séparé évolue en un état intriqué dès qu'il existe un couplage entre les variables d'espace et de spin, et aussi que la séparation espace - spin n'est pas un invariant de Lorentz.

Bien sûr, toutes les observables n'impliquent pas le spin, loin s'en faut. Par exemple, une coordonnée cartésienne est, en représentation-q, une simple multiplication, qui n'affecte pas le degré de spin. Pour un spin 1/2, en ordonnant les lignes et les colonnes toujours de la même façon, l'observable x (opérateur agissant dans l'espace des états $\mathcal{E}^{(\text{orb})}$) est ainsi représentée par la matrice 2×2 :

$$x = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} , \tag{20.45}$$

ce qui signifie que le vecteur résultant de l'action de x sur le spineur $|\Psi\rangle$ est :

$$x|\Psi\rangle \quad \longleftrightarrow \quad \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} \begin{bmatrix} \Psi_+(\vec{r}) \\ \Psi_-(\vec{r}) \end{bmatrix} = \begin{bmatrix} x\,\Psi_+(\vec{r}) \\ x\,\Psi_-(\vec{r}) \end{bmatrix} . \tag{20.46}$$

De tels opérateurs, qui ne touchent pas au spin, sont appelés opérateurs orbitaux. Le moment conjugué est de ceux-ci ; sa composante p_x sera ainsi représentée par la matrice :

$$p_x = -\,\mathrm{i}\hbar \begin{bmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial x} \end{bmatrix} , \tag{20.47}$$

avec la signification[29] :

$$p_x|\Psi\rangle \quad \longleftrightarrow \quad -\,\mathrm{i}\hbar \begin{bmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial x} \end{bmatrix} \begin{bmatrix} \Psi_+(\vec{r}) \\ \Psi_-(\vec{r}) \end{bmatrix} = -\,\mathrm{i}\hbar \begin{bmatrix} \partial_x\,\Psi_+(\vec{r}) \\ \partial_x\,\Psi_-(\vec{r}) \end{bmatrix} . \tag{20.48}$$

D'une façon générale, et par définition, les opérateurs orbitaux sont représentés dans l'espace des états de spin $|M_S\rangle$ par des matrices diagonales scalaires. En désignant par $\mathbf{1}_S$ l'identité dans cet espace, un opérateur orbital est toujours de la forme $f(\vec{r},\vec{p}) \otimes \mathbf{1}_S$, ou plus simplement $f(\vec{r},\vec{p})\,\mathbf{1}_S$; par exemple $p_x = -\mathrm{i}\hbar\partial_x\,\mathbf{1}_S = -\mathrm{i}\hbar\mathbf{1}_S\partial_x$; le plus souvent, on omet aussi $\mathbf{1}_S$, pour la simplicité de l'écriture.

Les opérateurs qui au contraire affectent seulement le spin sont appelés opérateurs de spin. Pour un spin 1/2, ils peuvent tous s'exprimer à l'aide des trois matrices de Pauli définies dans la section 18.4 ; en particulier, le moment cinétique de spin pour $S = 1/2$ s'exprime comme :

$$\boxed{\vec{S} = \frac{\hbar}{2}\,\vec{\sigma}} \tag{20.49}$$

Comme on l'a vu (voir (18.145)), toute fonction $f(\vec{S}\,)$ peut s'exprimer en combinaison linéaire de $\vec{S}$: $f(\vec{S}\,) = f_0\mathbf{1}_2 + \vec{f}.\vec{\sigma}$.

[29] $\partial_x \stackrel{\text{déf}}{=} \frac{\partial}{\partial x}$.

Il existe enfin des opérateurs mixtes, modifiant simultanément les degrés orbitaux et le spin. Soit par exemple le produit $L_z S_x$ agissant sur $\Psi_{M_S}(\vec{r})$; en coordonnées sphériques, on peut écrire :

$$L_z S_x \, \Psi_{M_S}(\vec{r}) = -\mathrm{i}\hbar \frac{\partial}{\partial \phi} \, S_x \, \Psi_{M_S}(\vec{r}) = -\mathrm{i}\hbar \; S_x \, \frac{\partial}{\partial \phi} \Psi_{M_S}(\vec{r}) \ . \tag{20.50}$$

Pour $S = 1/2$, compte tenu de (20.49) et de l'expression de la matrice de σ_x, la matrice représentant cet opérateur est sur la base $| + \frac{1}{2}\rangle$, $| - |\frac{1}{2}\rangle$ est :

$$L_z S_x = -\frac{\mathrm{i}}{2} \hbar^2 \begin{bmatrix} 0 & \frac{\partial}{\partial \phi} \\ \frac{\partial}{\partial \phi} & 0 \end{bmatrix} \ . \tag{20.51}$$

L'hermiticité de cette matrice ne saute pas aux yeux[30], bien que l'opérateur au premier membre est clairement hermitique puisque L_z et S_x le sont et commutent entre eux.

20.2 Magnétisme atomique : l'atome d'hydrogène

Il s'agit de mettre en évidence sur l'exemple le plus simple – l'atome d'hydrogène –, les sources fondamentales du magnétisme. En particulier, on va voir que dans son l'état fondamental – mais la conclusion est la même pour tous les états sphériques[31] de type S (i.e. $L = 0$) –, c'est le spin à lui seul qui explique le paramagnétisme observable de l'hydrogène. Plus généralement, le degré de liberté de spin est indispensable pour expliquer et décrire le magnétisme d'un atome complexe, notamment si ce dernier ne possède qu'un électron célibataire dans une couche s (c'est le cas de tous les alcalins). Si le spin n'existait pas, un tel atome à l'état fondamental serait dépourvu de moment magnétique permanent ; mettant à part la nature inattendue du spectre des impacts observés dans l'expérience de Stern et Gerlach (nombre discret de taches alors que classiquement, postulant que le moment magnétique *existe*, on prévoit une répartition continue), aucune déviation ne devrait de toute façon se produire.

Pour l'atome d'hydrogène, mettant de côté pour l'instant la structure fine[32] liée à V_{magn} défini en (20.20), et assimilant la masse réduite μ avec la masse m de l'électron, le Hamiltonien en l'absence de champ magnétique est :

$$H_0 = \frac{\vec{p}^{\,2}}{2m} + V(r) \ , \qquad V(r) = -\frac{e'^2}{r} \ ; \tag{20.53}$$

[30] La démonstration de l'hermiticité de (20.51) est laissée à titre d'exercice. Il suffit d'expliciter la condition d'hermiticité :

$$(L_z S_x |\Phi\rangle, |\Psi\rangle) = (|\Phi\rangle, L_z S_x |\Psi\rangle) \ . \tag{20.52}$$

en manipulant explicitement les deux composantes $\Phi_\pm$ et $\Psi_\pm$ des deux spineurs, et en faisant des intégrations par parties sur l'angle ϕ, toutes les fonctions de ϕ étant supposées périodiques et continues.

[31] Jusqu'à présent, on a désigné les états sphériques par la *minuscule* "s", tout comme la valeur du nombre quantique, l, donnant le carré du moment cinétique $\vec{L}$. L'usage est de désigner par des *capitales* ("S", L,...) les mêmes grandeurs relatives à un atome à plus d'un électron. Les arguments développés ici ayant une portée dépassant le cadre du plus simple des atomes, on s'autorise, même à son propos, un glissement de notations qui prendra réellement son sens pour les atomes complexes (voir chapitre 27).

[32] Le traitement complet en présence de structure fine est présenté dans la sous-section 23.3.2.

En l'absence d'une composante électrique ($\vec{\mathcal{E}} \equiv 0$) et pour un champ magnétique statique, le potentiel scalaire U peut être pris nul (voir Tome I, éq. (7.144)) ; $\vec{A}(\vec{r})$ désignant le (un) potentiel-vecteur, le Hamiltonien classique est (voir Tome I, éqs. (7.18) et (7.45), et problème 7.4.1) :

$$\boxed{H_{\text{classique}} = \frac{1}{2m}(\vec{p} - e\vec{A})^2 + V(r)} \tag{20.54}$$

où e désigne la charge de l'électron. Le Hamiltonien quantique en représentation-q s'obtient en remplaçant le vecteur $\vec{p}$ par $-\mathrm{i}\hbar\vec{\nabla}$. En respectant bien l'ordre des opérateurs ($\vec{A}(\vec{r})$ dépend de $\vec{r}$!), il vient ainsi :

$$H = \frac{\vec{p}^{\,2}}{2m} - \frac{e}{2m}(\vec{p}.\vec{A} + \vec{A}.\vec{p}) + \frac{e^2\vec{A}^{\,2}}{2m} + V(r) \equiv H_0 - \frac{e}{2m}(\vec{p}.\vec{A} + \vec{A}.\vec{p}) + \frac{e^2\vec{A}^{\,2}}{2m} \ . \tag{20.55}$$

En prenant l'axe Oz le long du champ magnétique, un choix possible de potentiel-vecteur[33] est le suivant :

$$\vec{A} = \frac{1}{2}\vec{\mathcal{B}} \times \vec{r} \quad \Longrightarrow \quad \vec{\mathcal{B}} = \vec{\nabla} \times \vec{A} = (0,\, 0,\, \mathcal{B}) \ . \tag{20.56}$$

Ce potentiel-vecteur est visiblement à divergence nulle : $\vec{\nabla}.\vec{A}(\vec{r}) = 0$; comme pour toute fonction vectorielle $\vec{f}(\vec{r})$, on a :

$$\boxed{[\vec{p},\, \vec{f}(\vec{r})] = -\mathrm{i}\hbar\vec{\nabla}.\vec{f}(\vec{r})} \tag{20.57}$$

le choix de la jauge (20.56) entraîne que l'opérateur $\vec{p}$ commute ici avec $\vec{A}(\vec{r})$. En outre, le report de l'expression (20.56) de $\vec{A}$ dans (20.55), suivi d'une permutation circulaire du produit mixte[34] $\vec{p}.(\vec{\mathcal{B}} \times \vec{r})$, fait apparaître le moment cinétique orbital $\vec{L} = \vec{r} \times \vec{p}$; on obtient ainsi :

$$H = H_0 - \frac{e}{2m}\,\vec{L}.\vec{\mathcal{B}} + \frac{e^2\vec{\mathcal{B}}^{\,2}}{8m}\,r^2\sin^2\theta \ , \tag{20.58}$$

θ étant l'angle entre $\vec{r}$ et $\vec{\mathcal{B}}$. Le terme linéaire en champ a la forme classique $-\vec{\mu}_{\text{orb}}.\vec{\mathcal{B}}$ avec $\vec{\mu}_{\text{orb}} = \frac{e}{2m}\vec{L}$ et a tout ce qu'il faut pour donner du paramagnétisme – pourvu que le moment cinétique orbital n'ait pas une moyenne nulle... Le terme quadratique[35] en champ, quant à lui, est de nature diamagnétique : il donne visiblement une variation positive d'énergie et est donc déstabilisant (le moment induit est en sens contraire du champ appliqué, conformément à la loi de Lenz). En fait, dans les conditions usuelles, et pour un élément léger, il est très petit devant le terme linéaire ; l'ordre de grandeur d'énergie du rapport de ces termes est :

$$\frac{E_{\text{dia}}}{E_{\text{para}}} \sim \frac{|e|\mathcal{B}}{4\hbar}\langle r^2\sin^2\theta\rangle \sim \frac{|e|\mathcal{B}}{4\hbar}\,a_0^2 \simeq 10^{-6}\,\mathcal{B}_{\text{Tesla}} \ ; \tag{20.59}$$

---

[33] Cette jauge est parfois appelée *jauge cylindrique.*

[34] Rappelons que, tout comme avec des vecteurs ordinaires, $r \times \vec{p} = -\vec{p} \times \vec{r}$, puisque $[u,\, p_v] = 0$ si $u \neq v$.

[35] Ce caractère à lui-seul montre que l'interaction ainsi représentée se produit entre le champ et un moment *induit* par le champ.

compte tenu de ceci, ce terme est négligé dans toute la suite[36].

L'expression (20.58), obtenue à partir du Hamiltonien classsique, ignore par construction un aspect essentiel : le magnétisme lié au spin. L'interaction correspondante doit être ajoutée "à la main" ; elle a l'expression usuelle $-\vec{\mu}_S.\vec{\mathcal{B}}$, $\vec{\mu}_S$ étant relié au spin $\vec{S}$ suivant (20.18). Cette adjonction étant faite, l'expression (20.58) (débarrassée du terme diamagnétique) devient[37] $H = H_0 + H_{\text{Zeeman}}$ où H_{Zeeman} est[38] :

$$\boxed{H_{\text{Zeeman}} = -\frac{e}{2m}\,(\vec{L} + g_e\vec{S}\,).\vec{\mathcal{B}} \equiv -\mu_B\hbar^{-1}(\vec{L} + g_e\vec{S}\,).\vec{\mathcal{B}} \qquad (g_e \simeq 2)} \tag{20.60}$$

Ce terme correctif est petit[39] devant H_0, dans les conditions usuelles. En effet, en ordre de grandeur et compte tenu du fait que $\langle\|\vec{L}\|\rangle \sim \langle\|\vec{S}\|\rangle \sim \hbar$, il vaut (voir (20.16)) :

$$\langle H_{\text{Zeeman}}\rangle \sim |\mu_B|\,\mathcal{B} \sim 10^{-5}\,\text{eV}\,\mathcal{B}_{\text{Tesla}} \ . \tag{20.61}$$

Comme les différences d'énergie pour les premiers niveaux atomiques sont de l'ordre de l'eV ou plus, il est clair que H_{Zeeman} est, pour les champs ordinaires, une petite correction. D'un autre côté, l'estimation faite en (20.8) montre que $\langle H_{\text{Zeeman}}\rangle$ est *a priori* comparable avec l'interaction spin-orbite, responsable de la structure fine. Pour éviter toute complication ici hors de propos – tout en gardant la cohérence –, on supposera le champ assez fort pour que le terme Zeeman domine largement les variations produisant cette structure fine[40], ce qui revient finalement à oublier cette dernière.

L'ensemble d'opérateurs (H, $\vec{L}^{\,2}$, L_z, $\vec{S}^{\,2}$, S_z) constitue visiblement un ensemble d'observables qui commutent. Les états propres de H_0, notés $|n, L, M_L, M_S\rangle$, forment une base sur laquelle on peut donc d'emblée écrire :

$$\langle n, L, M_L, M_S|H_{\text{Zeeman}}|n, L', M'_{L'}, M'_S\rangle \propto \delta_{LL'}\,\delta_{M_L M'_{L'}}\,\delta_{M_S M'_S} \ , \tag{20.62}$$

les symboles de Kronecker venant du fait que H_{Zeeman} commute avec chacun des opérateurs $\vec{L}^{\,2}$, L_z, $\vec{S}^{\,2}$ et S_z (pour un atome à *un* électron, si L varie d'un état à l'autre, S ne peut prendre qu'une seule valeur, $S = 1/2$). D'un autre côté, rien ne permet

[36]Évidemment, si le terme linéaire – qui donne la contribution paramagnétique – est de moyenne nulle (exemple : atome dans l'état singulet de spin (si nombre pair d'électrons) et dans un état S ($L = 0$) pour le moment cinétique orbital), le terme diamagnétique à lui seul représente tout l'effet du champ magnétique, du moins tant que l'on se borne à un formalisme de perturbation. On sait que le diamagnétisme est toujours présent – au contraire du paramagnétisme – et qu'il produit des effets très faibles.

[37]La limite de l'équation relativiste de Dirac produit tout naturellement le terme ici ajouté à la main et donne le Hamiltonien dit de Pauli (voir section 20.7).

[38]C'est ce terme qui est responsable de l'effet Zeeman (normal ou anormal), d'où la notation employée.

[39]Cette expression signifie que les valeurs moyennes typiques de H_{Zeeman} sont petites devant les *différences* des énergies propres fournies par H_0. La pertinence de ces différences d'énergie deviendra claire lors de l'exposé de la théorie des perturbations (voir chapitre 23, section 23.2, éq. (23.103)).

[40]C'est la situation dite de l'effet Paschen - Back. Dans le cas opposé (champ très faible en présence de structure fine, effet Zeeman strict), la bonne base est celle du moment cinétique total $\vec{J}$, soit $|nJM_JLS\rangle$. Chaque niveau de structure fine éclate en plusieurs niveaux très proches, l'écart étant toujours proportionnel à $\mathcal{B}$. En tant que combinaison linéaire d'opérateurs vectoriels, $\vec{\mu}_{\text{at}}$ est un opérateur vectoriel vis-à-vis de $\vec{J} = \vec{L} + \vec{S}$, dont le théorème de Wigner - Eckart (section 18.6) permet de calculer la valeur moyenne dans un sous-espace n, J donnés.

d'affirmer que les éléments de matrice de H_{Zeeman} entre deux états ne différant que par le nombre quantique n sont nuls – d'ailleurs, il ne le sont pas en général. Il en résulte que la résolution exacte de l'équation aux vecteurs propres de H_{Zeeman} exige *a priori* la considération simultanée de tous les états $\{|n, L, M_L, M_S\rangle\}_n$, L, M_L et M_S pouvant être fixés. En pratique, compte tenu de la petitesse de H_{Zeeman} par rapport aux différences d'énergie des niveaux de H_0, on résout cette question par la technique de perturbation, qui sera exposée en détail dans le chapitre 23. Ici, la question est juste de mettre en évidence simplement la valeur non-nulle du moment magnétique atomique, et on se doute qu'elle est donnée, à une très bonne approximation[41], par les valeurs moyennes du facteur de $-\vec{\mathcal{B}}$ dans H_{Zeeman}, calculées sur les états propres de H_0. Ces valeurs moyennes sont en fait le moment magnétique de l'atome dans la limite du champ nul (toujours en omettant la structure fine) et représentent donc sa valeur intrinsèque, en-dehors de toute contribution induite. Admettant ceci, la valeur moyenne du moment magnétique atomique est :

$$\langle \vec{\mu}_{\text{at}} \rangle_{n,L,M_L,M_S} = \langle n, L, M_L, M_S | \frac{\mu_{\text{B}}}{\hbar} (\vec{L} + g_{\text{e}} \vec{S}) | n, L, M_L, M_S \rangle \ . \tag{20.63}$$

Dans un état propre de $(\vec{S}^{\,2}, \, S_z)$, les valeurs moyennes de S_x et de S_y sont nulles, d'où :

$$\langle \vec{\mu}_{\text{at}} \rangle_{n,L,M_L,M_S} = \mu_{\text{B}} \, (M_L + g_{\text{e}} M_S) \; (0, \, 0, \, 1) \ . \tag{20.64}$$

Ainsi, quand Oz est choisi comme axe de quantification, seule la composante μ_z est non-nulle en moyenne et vaut :

$$\langle \mu_{\text{at}, \, z} \rangle_{n,L,M_L,M_S} = \mu_{\text{B}} \, (M_L + g_{\text{e}} M_S) \equiv \frac{e\hbar}{2m} \; (M_L + g_{\text{e}} M_S) \ . \tag{20.65}$$

Une telle moyenne est *toujours* différente de zéro : même dans un état S ($L = 0$) – le fondamental par exemple –, pour lequel M_L est forcément nul, il reste le nombre M_S qui prend lui les deux seules valeurs $M_S = \pm 1/2$: quel que soit son état, l'atome d'hydrogène possède donc toujours – grâce au spin de l'électron – un moment magnétique non-nul ; sa composante suivant Oz est bornée inférieurement en module par $(g_{\text{e}}/2) \, |e| \hbar/(2m)$:

$$\boxed{|\langle \mu_{\text{at}, \, z} \rangle| \geq \frac{1}{2} \, g_{\text{e}} \frac{|e|\hbar}{2m} \simeq \frac{|e|\hbar}{2m} \equiv |\mu_{\text{B}}|} \tag{20.66}$$

La même conclusion vaut pour les alcalins et, plus généralement, pour les atomes à nombre *impair* d'électrons. Le magnéton de Bohr, $|\mu_{\text{B}}|$, ressort bien comme l'unité fondamentale de moment magnétique atomique. Plus important : au moins pour un atome à un électron[42], le spin fournit d'une part la source fondamentale microscopique du magnétisme, d'autre part et parce qu'il relève de la *quantification*, il rend caduc le théorème de Miss van Leeuwen (voir Tome I, p. 359).

[41] Plus précisément à des corrections quadratiques en $\mathcal{B}$ près, venant de la correction au premier ordre des états propres ; ces corrections quadratiques représentent une part de l'effet inductif global.

[42] On verra plus tard (notamment dans le chapitre 27) comment le magnétisme d'un atome complexe résulte d'une coopération subtile entre les forces de Coulomb (purement *électrostatiques* !), le spin et l'indiscernabilité des particules identiques en Mécanique quantique.

20.3 Rotation d'un spin

Rappelons que le groupe des rotations dans $\mathbb{R}^3$, SO(3), joue un rôle fondamental en Physique puisque l'isotropie de l'espace, admise d'emblée par principe, impose aux lois physiques, et donc à toute prévision théorique, d'être invariantes dans toute opération de ce groupe. Comme on l'a vu (voir chapitre 17), toute rotation R dans l'espace physique induit une certaine transformation $\mathcal{R}$ dans l'espace des états $\mathcal{E}$ qui, pour les rotations, est forcément unitaire. Pour la rotation d'un angle θ autour de l'axe défini par le vecteur $\vec{n}$ de norme égale à 1, l'opérateur linéaire unitaire est fabriqué avec le moment cinétique $\vec{J}$ (voir éq. (17.98)) :

$$\mathcal{R}(\theta,\, \vec{n}) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\,\theta\,\vec{J}.\vec{n}} \quad . \tag{20.67}$$

Pour un atome, quand on tient compte du spin, $\vec{J}$ est la somme donnée en (20.10), pour les raisons physiques qui ont permis d'argumenter en partant d'une évidence expérimentale – l'existence du moment magnétique de l'électron – pour en arriver à la conclusion inévitable de l'existence du *spin* de l'électron. C'est bien vis-à-vis du moment total $\vec{J}$ que les vecteurs $\vec{r}$, $\vec{p}$, $\vec{L}$ et $\vec{S}$ satisfont les relations caractéristiques (17.117), requises pour qu'ils soient des opérateurs vectoriels.

Autant le mode de transformation par rotation d'un moment cinétique orbital est clair, puisque ce dernier résulte d'un *mouvement* dont on peut se faire une représentation imagée et en tout cas s'exprime en $\vec{r} \times \vec{p}$, autant celui d'un moment cinétique de spin est loin d'être évident[43]. Pour les mêmes raisons, la transformation des états de spin n'est pas intuitive. Afin de déterminer ce mode de transformation, on s'appuie sur le fait que $\vec{\mu}_{\mathrm{at}}$ doit être, physiquement, un opérateur vectoriel ; de ce fait, les valeurs moyennes de ses composantes se comportent par rotation comme celles d'un vecteur de la géométrie ordinaire. Alors, l'opérateur de rotation dans l'espace du spin se construit avec $\vec{S}$ exactement comme la rotation dans $\mathbb{R}^3$ se construit avec $\vec{L}$; comme le moment cinétique total est maintenant $\vec{L} + \vec{S}$, la transformation $\mathcal{R}$ induite dans l'espace des états est toujours de la forme (20.67), avec précisément $\vec{J} = \vec{L} + \vec{S}$.

En vertu de ceci, de (20.67), de (20.10) et du fait que $[\vec{L},\, \vec{S}] = 0$, on a[44] :

$$\mathcal{R}(\theta,\, \vec{n}) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,(\vec{L}+\vec{S}).\vec{n}} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,\vec{L}.\vec{n}}\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,\vec{S}.\vec{n}} \equiv \mathcal{R}^{(\mathrm{orb})}(\theta,\, \vec{n}) \otimes \mathcal{R}^{(\mathrm{spin})}(\theta,\, \vec{n}) \quad . \tag{20.68}$$

Sous l'effet de la rotation R dans $\mathbb{R}^3$, un état *factorisé* $|\Psi\rangle = |\psi\rangle \otimes |\chi\rangle$ devient $|\Psi'\rangle$:

$$|\Psi'\rangle \stackrel{\text{déf}}{=} \mathcal{R}(\theta,\, \vec{n})|\Psi\rangle = \mathcal{R}^{(\mathrm{orb})}(\theta,\, \vec{n})|\psi\rangle \otimes \mathcal{R}^{(\mathrm{spin})}(\theta,\, \vec{n})|\chi\rangle \tag{20.69}$$

montrant que l'état de spin est bel et bien affecté par la rotation. Comme la transformation de la partie orbitale a déjà été élucidée au chapitre 18, on examine dans la suite la seule transformation dans l'espace de spin.

[43] Concernant le comportement d'un spin par réflexion, voir [4], chapitre 12, section 8.

[44] C'est la commutation de $\vec{L}$ et $\vec{S}$ qui permet d'écrire que l'exponentielle de la somme est égale au produit des exponentielles.

Pour simplifier, on fixe la valeur du spin à $S = 1/2$ – c'est aussi le cas le plus important en pratique. Dans ces conditions, on a précisément (voir (20.49)) :

$$\mathcal{R}^{(\text{spin }1/2)}(\theta,\, \vec{n}) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\frac{\hbar}{2}\vec{n}.\vec{\sigma}} \quad . \tag{20.70}$$

En développant l'exponentielle en série et en utilisant $\sigma_u^2 = \mathbf{1}_2$, on voit que, pour un spin 1/2 :

$$\boxed{\mathcal{R}^{(\text{spin }1/2)}(\theta,\, \vec{n}) = \cos\frac{\theta}{2}\,\mathbf{1}_2 \;-\; \mathrm{i}\,\vec{n}.\vec{\sigma}\,\sin\frac{\theta}{2} \qquad (S = 1/2)} \tag{20.71}$$

Il est facile d'écrire la matrice de 2×2 représentant cet opérateur sur la base propre de S_z ; en utilisant les expressions des matrices de Pauli (18.135), on obtient pour la base ordonnée suivant $|+\frac{1}{2}\rangle$, $|-\frac{1}{2}\rangle$:

$$\mathcal{R}^{(\text{spin }1/2)}(\theta,\, \vec{n}) = \begin{bmatrix} \cos\frac{\theta}{2} - \mathrm{i}n_z \sin\frac{\theta}{2} & -\mathrm{i}(n_x - \mathrm{i}n_y)\sin\frac{\theta}{2} \\ -\mathrm{i}(n_x + \mathrm{i}n_y)\sin\frac{\theta}{2} & \cos\frac{\theta}{2} + \mathrm{i}n_z \sin\frac{\theta}{2} \end{bmatrix} \quad . \tag{20.72}$$

L'ensemble des matrices $\mathcal{R}^{(\text{spin }1/2)}(\theta, \vec{n})$ a manifestement une structure de groupe[45], appelé groupe[46] SU(2) ; ces matrices, agissant sur les spineurs (deux composantes), constituent une *représentation* de dimension 2 du groupe des rotations, dont les générateurs sont les composantes du moment cinétique $\vec{J} = \frac{\hbar}{2}\vec{\sigma}$; noter que, à la différence des matrices de SO(3), celles de SU(2) sont complexes[47].

La présence de l'angle moitié $\theta/2$ est visiblement liée à la valeur demi-entière du spin et a une conséquence un peu surprenante. En effet, si on prend $\theta = 2\pi$ (un tour complet), on s'attendrait à ce que la rotation soit la transformation identité. Il n'en est rien ; les formules précédentes donnent dans ce cas :

$$\mathcal{R}^{(\text{spin }1/2)}(\theta = 2\pi, \vec{n}) = \cos\pi\,\mathbf{1}_2 = -\,\mathbf{1}_2 \quad . \tag{20.74}$$

[45]En particulier, le sous-ensemble des rotations autour d'un axe donné forme un groupe commutatif (abélien) :

$$\mathcal{R}^{(\text{spin }1/2)}(\theta_1,\, \vec{n})\mathcal{R}^{(\text{spin }1/2)}(\theta_2,\, \vec{n}) = \mathcal{R}^{(\text{spin }1/2)}(\theta_1 + \theta_2,\, \vec{n}) \quad . \tag{20.73}$$

[46]"SU" pour *special unitary*.

[47]Il existe une relation précise entre SO(3) et SU(2) – voir problème 20.8.2. La correspondance entre une matrice hermitique de trace nulle A (voir (18.148) avec $a_0 = 0$) et un certain vecteur $\vec{a} \in \mathbb{R}^3$ permet de définir une application de SU(2) dans SO(3) faisant correspondre une certaine matrice $R_U \in$ SO(3) à toute matrice $U \in$ SU(2) (pour plus de détails, voir [12]). C'est en ce sens que SO(3) constitue une *représentation* de SU(2), mais la réciproque est fausse car les deux matrices U et $-U$ ont la même image $R_U = R_{-U}$. L'application considérée SU(2) $\rightarrow$ SO(3) est donc un homomorphisme, non un isomorphisme – sauf à introduire le groupe quotient SU(2)/Z_2 où R_U et R_{-U} sont réunies dans la même classe d'équivalence.

Il n'est pas possible de construire un isomorphisme, car SU(2) et SO(3) n'ont pas la même topologie : le premier est simplement connexe, le second est doublement connexe. Cette distinction peut s'apprécier en sachant que l'espace des paramètres de SO(3) est la boule de $\mathbb{R}^3$ de rayon π ; deux points diamétralement opposés de sa surface doivent être identifiés puisque la rotation de π autour de l'axe $\vec{n}$ est identique à la rotation de π autour de l'axe $-\vec{n}$: il existe donc des boucles non homotopes à zéro. Au contraire, l'espace des paramètres de SU(2) est la sphère unité plongée dans $\mathbb{R}^4$: sur la sphère, toute boucle peut être continûment réduite à un point.

Ainsi, quand, par exemple, l'appareil de mesure fait un tour complet, le spineur associé à $S = 1/2$ *change de signe* et l'état total $|\Psi\rangle = |\psi\rangle \otimes |\chi\rangle$ en fait autant. Quoique surprenant, ceci n'est pas gênant : les deux états $\pm |\psi\rangle \otimes |\chi\rangle$ ne diffèrent que par une phase globale et représentent en fait le même état physique. La description de l'état physique est donc inaltérée par la rotation d'un tour complet *tant qu'un seul état est impliqué* (mais voir Remarque 4 ci-dessous) ; en particulier, les valeurs moyennes d'observables sont inchangées puisqu'elles dépendent *quadratiquement* du vecteur d'état.

♦ *Remarques*

1. Le changement de signe d'un spineur par une rotation de 2π effectuée dans l'espace physique $\mathbb{R}^3$, alors qu'un vecteur se retrouve à l'identique, est la motivation fondamentale de l'introduction nécessaire de la notion de spineur dès qu'un espace vectoriel sur le corps des complexes est inévitable – et c'est le cas pour la Mécanique quantique. C'est pourquoi il est exact de dire que la notion de spin est *essentiellement quantique*. En Physique classique, toutes les quantités fondamentales sont des quantités réelles (positions, moments conjugués,...). Il en résulte que toutes les représentations du groupe des rotations seront des matrices dont tous les éléments sont des nombres réels. Pour des raisons subtiles liés à la *connexité* [12] des groupes SO(3) et SU(2), les représentations réelles de SO(3) sont toutes de dimension *impaire* ; $2J+1$ étant impair, seules les valeurs *entières* de J sont possibles et toutes les rotations de 2π donnent $\mathcal{R} = \mathbf{1}$: la Physique classique ne connaît que les moments cinétiques entiers parce qu'elle n'a pas besoin des autres.

2. La bizarrerie exprimée par (20.74) peut être énoncée sous forme plus générale ([68], § 7-6). En effet, pour un moment cinétique $\vec{J}$ quelconque, d'états propres $|j\,m\rangle$, on a :

$$\boxed{\mathcal{R}(\theta = 2\pi, \vec{n})|j\,m\rangle = (-1)^{2j}\,|j\,m\rangle \qquad \forall\, j,\, m} \tag{20.75}$$

Pour démontrer cette relation, commençons par prendre une rotation autour de Oz ; alors, on a :

$$R^{(j)}(\theta, \mathrm{O}z)|j\,m\rangle = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta J_z}\,|j\,m\rangle = \mathrm{e}^{-\mathrm{i}m\theta}\,|j\,m\rangle \ , \tag{20.76}$$

puisque le vecteur est propre de J_z, d'où :

$$R^{(j)}(2\pi, \vec{n})|j\,m\rangle = \mathrm{e}^{-2\mathrm{i}m\pi}\,|j\,m\rangle \ . \tag{20.77}$$

Toutes les valeurs de m ne diffèrant entre elles que par un entier relatif, l'égalité suivante est aussi vraie :

$$\mathrm{e}^{-2\mathrm{i}m\pi} = \mathrm{e}^{-2\mathrm{i}j\pi} = (-1)^{2j} \qquad \forall m \in \{-j,\, -j+1, \ldots,\, j\} \ . \tag{20.78}$$

Pour un autre axe de rotation, le résultat de l'action de $R^{(j)}(2\pi, \vec{n})$ sur $|j\,m\rangle$ est une combinaison linéaire des seuls $(2j+1)$ vecteurs j fixé, m variable, $\{|j\,m\rangle\}_{-j\leq m\leq +j}$, puisque $\vec{n}.\vec{J}$ commute avec $\vec{J}^2$ quel que soit $\vec{n}$; mais la même phase $(-1)^{2j}$ apparaît pour chacun d'entre eux, et se met donc en facteur – ce qui achève d'établir (20.75).

En conséquence, si j est entier, la rotation de 2π est bien l'identité ; dans le cas contraire :

$$R^{(j)}(2\pi, \vec{n})|j\,m\rangle = -\,|j\,m\rangle \qquad \text{(pour tout } j \text{ demi-entier)} \;; \tag{20.79}$$

ceci généralise (20.74) pour j demi-entier quelconque.

Notons que si le vecteur d'état peut éventuellement être affecté d'un signe lors d'une rotation de 2π, il n'en va pas de même pour une observable transformée ; $A' = RAR^\dagger$ étant une expression quadratique en R, tout signe finit par disparaître. Pour toute observable, on a donc :

$$[A,\, R^{(j)}(2\pi, \vec{n})] = 0 \quad \Longleftrightarrow \quad A\,R^{(j)}(2\pi, \vec{n}) = R^{(j)}(2\pi, \vec{n})\,A \;. \tag{20.80}$$

En particulier, la valeur moyenne de toute observable est invariante dans une rotation d'un tour complet.

3. Si l'on considère *tous* les vecteurs de type $|j\,m\rangle$, $\forall j$, m, on voit que l'espace vectoriel qu'ils engendrent peut être structuré en deux sous-espaces[48], l'un contenant tous les vecteurs ayant j entier, l'autre tous les vecteurs ayant j demi-entier[49]. Pour la simplicité, on désigne par $|\natural\rangle$ un vecteur quelconque du premier sous-espace (alors noté $\mathcal{E}_\natural$), et $|\flat\rangle$ un vecteur quelconque du second sous-espace noté $\mathcal{E}_\flat$; de la sorte, on a :

$$R^{(j)}(2\pi, \vec{n})|\natural\rangle = +\,|\natural\rangle \;, \qquad R^{(j)}(2\pi, \vec{n})|\flat\rangle = -\,|\flat\rangle \;. \tag{20.81}$$

Les deux sous-espaces $\mathcal{E}_\natural$ et $\mathcal{E}_\flat$ s'ignorent complètement au sens où aucune observable ne permet de passer de l'un à l'autre. En effet, d'après (20.80), toute observable A est insensible à une rotation d'un tour complet. Prenons l'élément de matrice de (20.80) entre deux vecteurs $|\natural\rangle$ et $|\flat\rangle$; il vient :

$$\langle\flat|AR^{(j)}(2\pi, \vec{n})|\natural\rangle = \langle\flat|R^{(j)}(2\pi, \vec{n})A|\natural\rangle \quad \Longleftrightarrow \quad \langle\flat|A|\natural\rangle = -\langle\flat|A|\natural\rangle \;, \tag{20.82}$$

d'où résulte :

$$\langle\flat|A|\natural\rangle = 0 \qquad \forall\, A \;. \tag{20.83}$$

Ainsi, aucun processus ne peut transmuter un état $|\natural\rangle$ en un état $|\flat\rangle$ et inversement, ceci constituant un exemple de règle de *supersélection*[50]. Il est ainsi impossible passer d'un état à nombre impair de fermions en un autre contenant un nombre pair de fermions ; un fermion ne peut se transformer en boson, et inversement. C'est aussi pourquoi la fission du neutron en un proton et un électron (rayonnement β^-) exige la participation d'au moins un autre partenaire dans les produits finals (en l'occurrence un antineutrino).

---

[48] On dit parfois *secteurs*.

[49] Ces deux sous-espaces sont mutuellement orthogonaux : en tant que vecteurs propres d'observables associés à des valeurs propres distinctes, tout vecteur $|j\,m\rangle$ est orthogonal à tout vecteur $|j'\,m'\rangle$ si $j \neq j'$ (et/ou) $m \neq m'$.

[50] Un autre exemple, très familier, d'une telle règle est fourni par la conservation de l'énergie : une transition qui viole la conservation de l'énergie est *strictement* interdite.

Une telle règle a pour vertu de rendre inessentielle une phase non-globale qui, dans un autre contexte, ne peut ni ne doit être délaissée[51]. En effet, soit un vecteur du genre :

$$|\Psi\rangle = |\natural\rangle + \mathrm{e}^{\mathrm{i}\alpha}|\flat\rangle \ . \tag{20.84}$$

Compte tenu de (20.83), la valeur moyenne d'une observable dans l'état $|\Psi\rangle$ est :

$$\langle\Psi|A|\Psi\rangle = \langle\natural|A|\natural\rangle + \mathrm{e}^{-\mathrm{i}\alpha}\mathrm{e}^{\mathrm{i}\alpha}\langle\flat|A|\flat\rangle = \langle\natural|A|\natural\rangle + \langle\flat|A|\flat\rangle \ . \tag{20.85}$$

Cette valeur moyenne est indépendante de la phase α – ce qui montre que cette dernière ne peut recevoir *ici* le moindre sens physique et qu'elle peut être complètement ignorée.

4. À l'inverse, pour *un j* demi-entier *donné*, la phase $\mathrm{i}\pi$ donnant le signe $-$ dans (20.79) est visible dans des expériences d'interférométrie, comme l'ont proposé [69] puis montré [70] Overhauser et al. dans une expérience couplant de façon remarquable la Mécanique quantique et la gravitation (plus précisément la pesanteur terrestre) ; presque simultanément, Rauch et al. [71] ont également mis en évidence le signe $-$ par interférométrie en faisant passer l'un des deux faisceaux de neutrons dans un champ magnétique. ♦

20.4 Retour sur le renversement du temps

On a vu au chapitre 17 que, pour une particule décrite par une seule fonction d'onde (particule sans spin), l'opération de renversement du temps, alors désignée par K, est telle que :

$$\Psi_{\text{renv}}(\vec{r},\, t) \stackrel{\text{déf}}{=} K\Psi(\vec{r},\, t) = \Psi^*(\vec{r}, -t) \ . \tag{20.86}$$

Dans ce contexte, pour toute fonction indépendante du temps, l'opération de renversement du temps est donc strictement identique à la conjugaison complexe ; dès lors, si on désigne par K_0 l'opération définie comme :

$$K_0\psi(\vec{r}) \stackrel{\text{déf}}{=} \psi^*(\vec{r}) \qquad \forall\psi \tag{20.87}$$

on a simplement :

$$\boxed{K = K_0 \qquad \text{(particule sans spin)}} \tag{20.88}$$

et, de toute évidence[52] :

$$K^2 \equiv K_0^2 = \mathbf{1} \ . \tag{20.89}$$

H lui-même est invariant par renversement du temps quand seules des puissances paires du moment conjugué y apparaissent et ceci assure que H, en représentation-q, est

[51] On sait bien que, en général, une phase dans un terme d'une combinaison linéaire est essentielle par le rôle qu'elle joue dans les interférences quantiques.

[52] (20.89) entraine aussi que $K_0^\dagger = K_0$.

représenté par un opérateur différentiel *réel*. Comme on l'a vu (voir (17.182) et (17.185)), la transformation des observables fondamentales est :

$$\vec{r}\,' = K\,\vec{r}\,K^\dagger = \vec{r}\ , \qquad \vec{p}\,' = K\,\vec{p}\,K^\dagger = -\vec{p}\ . \tag{20.90}$$

Comme déjà mentionné, les relations de commutation fondamentales – qui contiennent le scalaire imaginaire pur i – changent de signe par renversement du temps, rappelant au passage que K est un opérateur antiunitaire[53]. Lorsque la représentation est réelle, la coordonnée est représentée par une matrice réelle, le moment conjugué est représenté par une matrice imaginaire pure : en pareil cas, K coïncide strictement avec l'opération de conjugaison complexe effectuée cette fois sur les *matrices* ; cette équivalence opératoire n'est pas vraie en général[54].

De la transformation de $\vec{r}$ et $\vec{p}$ (20.90), il résulte immédiatement que le moment cinétique orbital est lui aussi changé en son opposé (voir (17.186)). Physiquement, ceci signifie simplement que la particule "renversée" tourne dans le sens opposé. En revanche, la transformation du moment cinétique de spin est loin d'être évidente, puisque l'on n'en possède pas une définition "fondamentale", en termes de grandeurs dynamiques. Cependant, comme il s'agit précisément d'un moment cinétique, on *postule* :

$$\boxed{K\,\vec{S}\,K^\dagger = -\vec{S}} \tag{20.92}$$

Ceci assure, notamment, que le moment cinétique total $\vec{J} = \vec{L} + \vec{L}$ est changé en son opposé par renversement du temps[55] :

$$K\,\vec{J}K^\dagger = -\vec{J} \iff K\,J_z\,K^\dagger = -J_z\ , \quad K\,J_\pm\,K^\dagger = -J_\mp\ . \tag{20.93}$$

Tout ceci est cohérent : le moment cinétique total ne saurait se transformer autrement, en vertu de l'argument physique qui a conduit de l'évidence expérimentale d'un moment magnétique propre de l'électron – impliquant nécessairement un moment cinétique intrinsèque – à la nécessaire conservation du moment angulaire total d'un système isolé. En outre, K commute visiblement avec toutes les transformations spatiales (rotations, translations, etc.) : toucher au temps ne saurait affecter les coordonnées spatiales ; toutefois, compte tenu du caractère particulier de K, il n'est pas inutile de vérifier la cohérence de ce qui précède. $\vec{J}$ étant le générateur des rotations, on a – précisément parce que K est antiunitaire :

$$K\,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,\vec{n}.\vec{J}}K^\dagger = \mathrm{e}^{-\frac{1}{\mathrm{i}\hbar}\theta\,\vec{n}.(-\vec{J})}\,KK^\dagger = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\theta\,\vec{n}.\vec{J}} \equiv \mathcal{R}_{\vec{n},\theta} \iff [K, \mathcal{R}_{\vec{n},\theta}] = 0\ . \tag{20.94}$$

La même vérification peut être faite avec $\vec{p}$, générateur des translations.

[53]La deuxième relation dans (20.90) montre que l'anticommutateur de K et du moment conjugué est nul :

$$K\vec{p}K^\dagger = -\vec{p} \iff K\vec{p} + \vec{p}K \equiv \{K, \vec{p}\} = 0\ . \tag{20.91}$$

[54]Tout dépend de la représentation choisie. Soit par exemple les trois matrices 3×3 d'un moment angulaire $J = 1$ (voir (18.156) et (18.160)). J_x et J_z ont des matrices réelles, cependant que celle de J_y est imaginaire pure. Il est bien clair que conjuguer ces matrices ne change pas $\vec{J}$ en $-\vec{J}$. La même remarque vaut dans le cas $S = 1/2$, essentiellement représenté par les trois matrices de Pauli.

[55]Se souvenir que si Ω est antiunitaire : $\forall$ scalaire λ, $\Omega\,\lambda = \lambda^*\,\Omega$, d'où l'égalité de droite dans (20.93).

Il s'agit maintenant de construire l'opérateur de renversement du temps pour une particule avec spin, que j soit entier ou demi-entier[56]. Dans la suite, on se met dans une représentation dite *standard*, dont la base est l'ensemble des vecteurs $|\alpha\, j\, m\rangle$ propres d'un ECOC[57] contenant d'une part des observables invariantes par renversement du temps, dont les nombres quantiques sont collectivement dénotés α, d'autre part le couple $(\vec{J}^2,\, J_z)$. K change le signe de chaque composante de $\vec{J}$, mais laisse $\vec{J}^2$ invariant, en conséquence de $[K,\, \vec{J}^2] = 0$; prenant l'élément de matrice entre deux états ne différant que par j, on a :

$$\langle \alpha\, j\, m | K \vec{J}^2 | \alpha\, j'\, m\rangle = \langle \alpha\, j\, m | \vec{J}^2 K | \alpha\, j'\, m\rangle \ , \qquad (20.95)$$

soit :

$$[j(j+1) - j'(j'+1)]\, \langle \alpha\, j\, m | K | \alpha\, j'\, m\rangle = 0 \ . \qquad (20.96)$$

K n'ayant donc pas d'élément de matrice non-nul entre deux vecteurs de j différents, on peut fixer une fois pour toute la valeur de j ; le même argument vaut pour les observables faisant apparaître les nombres α : on peut donc tout autant fixer α. Ce faisant, on se trouve dans le sous-espace de dimension $2j+1$ engendré par les vecteurs ayant tous les mêmes j et α et différant seulement par leur m. Pour simplifier les notations, on pose désormais $|\alpha\, j\, m\rangle \equiv |m\rangle$.

L'opérateur K sera complètement défini quand on aura trouvé le résultat de l'action de K sur ces vecteurs. Physiquement, renverser le temps inverse les moments cinétiques – c'est bien ce qu'exprime (20.93) ; on s'attend donc à ce que $K|m\rangle \propto |-m\rangle$, ce qui assure déjà que la valeur propre de J_z change bien de signe, et autorise seulement un facteur de phase dépendant *a priori* de m :

$$K|m\rangle = \mathrm{e}^{\mathrm{i}\lambda(m)}\, |-m\rangle \ , \qquad (20.97)$$

l'objectif étant de trouver la fonction $\lambda(m)$. (20.97) définit *ipso facto* $K^\dagger$; en effet, partant de la définition (17.32) de l'adjoint d'un opérateur antilinéaire, on a :

$$(K^\dagger|m\rangle,\, |m'\rangle) = (|m\rangle,\, K|m'\rangle)^* = (|m\rangle,\, \mathrm{e}^{\mathrm{i}\lambda(m')}|-m'\rangle)^* = \\ \mathrm{e}^{-\mathrm{i}\lambda(m')}(|m\rangle,\, |-m'\rangle)^* = \mathrm{e}^{-\mathrm{i}\lambda(m')}\,\delta_{m\,-m'} \ ; \qquad (20.98)$$

prenant les complexes conjugués :

$$(|m'\rangle,\, K^\dagger|m\rangle) = \mathrm{e}^{\mathrm{i}\lambda(-m)}\,\delta_{m\,-m'} \quad \Longleftrightarrow \quad K^\dagger|m\rangle = \mathrm{e}^{\mathrm{i}\lambda(-m)}\,|-m\rangle \ . \qquad (20.99)$$

Explicitons maintenant le premier membre de l'équation (20.93) :

$$KJ_zK^\dagger|m\rangle = KJ_z\mathrm{e}^{\mathrm{i}\lambda(-m)}\,|-m\rangle = -m\hbar K\mathrm{e}^{\mathrm{i}\lambda(-m)}\,|-m\rangle \ ; \qquad (20.100)$$

comme K est antilinéaire, cette dernière expression est $-m\hbar\,\mathrm{e}^{-\mathrm{i}\lambda(-m)}K\,|-m\rangle$, d'où, utilisant (20.97) avec $-m$:

$$KJ_zK^\dagger|m\rangle = -m\hbar\,\mathrm{e}^{-\mathrm{i}\lambda(-m)}\mathrm{e}^{\mathrm{i}\lambda(-m)}|+m\rangle = -m\hbar|m\rangle \equiv -J_z|m\rangle \ , \qquad (20.101)$$

[56]Pour une autre présentation de la suite, voir [15], XV-18, p. 571.

[57]Le Hamiltonien n'est pas forcément dans l'ECOC ; l'argument qui suit vaut donc tout autant que le système ait ou n'ait pas la symétrie sphérique.

ce qui montre que, comme attendu, J_z est transformé en $-J_z$; maintenant :

$$KJ_+K^\dagger|m\rangle = KJ_+\mathrm{e}^{\mathrm{i}\lambda(-m)}|-m\rangle = \hbar\mathrm{e}^{-\mathrm{i}\lambda(-m)}\sqrt{j(j+1)-m(m-1)}\,K|-m+1\rangle = \hbar\mathrm{e}^{-\mathrm{i}\lambda(-m)}\sqrt{j(j+1)-m(m-1)}\,\mathrm{e}^{\mathrm{i}\lambda(-m+1)}|m-1\rangle \ , \quad (20.102)$$

où l'égalité (20.97) avec $-m+1$ été utilisée. Conformément à (20.93), ceci doit être strictement égal à $-J_-|m\rangle$, d'où quel que soit m :

$$-\lambda(-m)+\lambda(-m+1)=\pi \quad (2\pi) \quad \Longleftrightarrow \quad \lambda(m+1)-\lambda(m)=\pi \quad (2\pi) \ . \quad (20.103)$$

La dernière condition $KJ_-K^\dagger = -J_+$ fournit la même équation. Finalement, – que j soit entier ou demi-entier – le passage d'une valeur de m à la suivante fait augmenter de π la phase λ ; l'origine des phases étant inessentielle, on peut donc prendre :

$$\lambda(m) = m\pi \qquad (2\pi) \ , \quad (20.104)$$

ce qui achève de définir l'opérateur de renversement du temps pour un moment cinétique quelconque ; revenant à la notation explicite $|m\rangle \equiv |\alpha\, j\, m\rangle$:

$$K|\alpha\, j\, m\rangle = \mathrm{e}^{\mathrm{i}m\pi}\,|\alpha\, j\ -m\rangle \ , \quad (20.105)$$

et comme la base est propre de J_z, ceci s'énonce encore :

$$\boxed{K|\alpha\, j\, m\rangle = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\pi J_z}\,|\alpha\, j\ -m\rangle} \quad (20.106)$$

Formellement, et avec ce choix d'origine pour la phase, l'action de K sur $|\alpha\, j\, m\rangle$ est la même que celle d'une rotation de $\pi\,(2\pi)$ autour de l'axe Oz effectuée sur $|\alpha\, j\ -m\rangle$. Ces résultats sont vrais $\forall\ j$ et en particulier pour j entier auquel cas il peut s'agir d'un moment cinétique orbital. Alors, la représentation-q des $|\alpha\, j\, m\rangle$ fait apparaître les $Y_{l\,m}$ et (20.105) se transcrit comme suit :

$$KY_{l\,m} = \mathrm{e}^{\mathrm{i}m\pi}\,Y_{l\,-m} = (-1)^m\,Y_{l\,-m} \ . \quad (20.107)$$

Le second membre n'est autre que[58] $(Y_{lm})^*$: dans ce cas précis, l'opérateur K se réduit bien à la seule conjugaison complexe.

Cela étant acquis, que vaut le carré de K ? Par (20.105), on a :

$$K^2|m\rangle = KK|m\rangle = K(\mathrm{e}^{\mathrm{i}m\pi}|-m\rangle) = \mathrm{e}^{-\mathrm{i}m\pi}K|-m\rangle = \mathrm{e}^{-\mathrm{i}m\pi}\mathrm{e}^{-m\mathrm{i}\pi}|m\rangle = \mathrm{e}^{-2\mathrm{i}m\pi}|m\rangle \ , \quad (20.108)$$

ce que l'on peut aussi voir avec l'expression (20.106), compte tenu de la commutation (20.94). Comme, dans tous les cas, m ne diffère de j que par un entier, on a :

$$\mathrm{e}^{2\mathrm{i}m\pi} = \mathrm{e}^{2\mathrm{i}j\pi} \qquad \forall j,\ m=-j,\ -j+1,\ldots,+j \ . \quad (20.109)$$

[58] voir chapitre 18, éq. (18.88).

En définitive, sur cette base standard, le carré K^2 de K a la même action que $e^{2ij\pi}\,\mathbf{1}$, d'où finalement l'important résultat[59] :

$$\boxed{K^2 = \begin{cases} +\mathbf{1} & \text{si spin entier} \\ -\mathbf{1} & \text{si spin demi-entier} \end{cases}} \tag{20.110}$$

Par le théorème fondamental d'addition des moments cinétiques, le caractère entier ou demi-entier du spin *total* d'un système composite est lié au nombre de particules de spin demi-entier, $N_{1/2}$; si ce nombre est pair, le spin est entier, s'il est impair, le spin est demi-entier. En définitive :

$$\boxed{K^2 = (-1)^{N_{1/2}}\,\mathbf{1}} \ . \tag{20.111}$$

Ce résultat permet de démontrer l'existence de la dégénérescence de Kramers pour les systèmes à spin demi-entier, annoncé p. 700. En effet, dans le cas où $K^2 = -\mathbf{1}$, on a :

$$K = -K^{-1} = -K^\dagger \ , \tag{20.112}$$

d'où, en partant de la définition (17.32) de l'adjoint d'un opérateur antiunitaire :

$$(K^\dagger|\Psi\rangle,\,|\Psi\rangle) = (|\Psi\rangle,\,K|\Psi\rangle)^* = (K|\Psi\rangle,\,|\Psi\rangle) = -(K^\dagger|\Psi\rangle,\,|\Psi\rangle) \ . \tag{20.113}$$

Les termes extrêmes de cette chaîne montrent que le produit scalaire $(K^\dagger\Psi\rangle,\,|\Psi\rangle)$ est nul :

$$(K^\dagger|\Psi\rangle,\,|\Psi\rangle) = 0 \quad \Longleftrightarrow \quad (|\Psi\rangle,\,K|\Psi\rangle) = 0 \ . \tag{20.114}$$

Si H est invariant par renversement du temps, $|\Psi\rangle$ et $K|\Psi\rangle$ d'une part ont la même énergie et d'autre part sont orthogonaux : il existe donc en ce cas une dégénérescence forcément paire, c'est la dégénérescence de Kramers. Un champ magnétique brise l'invariance par renversement du temps et lève donc la dégénérescence de Kramers.

♦ *Remarque*

Il est possible de d'écrire K sous forme d'un produit d'opérateurs[60], comme on l'a fait partiellement pour écrire (20.106) ; il suffit pour cela d'introduire l'opérateur qui transforme $|\alpha\,j\,m\rangle$ en $|\alpha\,j\,-m\rangle$. Sur la base standard, cet opérateur est représenté par une matrice qui a des 1 partout sur l'antidiagonale[61] et des zéros partout ailleurs[62]. Visiblement, le carré de cet opérateur est l'identité $\mathbf{1}$: il est donc assez naturel de le noter[63] conventionnellement $\sqrt{\mathbf{1}}$. Ainsi, par définition :

$$\sqrt{\mathbf{1}}\,|\alpha\,j\,m\rangle = |\alpha\,j\,-m\rangle \ . \tag{20.115}$$

[59] Comme un moment demi-entier ne peut provenir *que* du spin, on peut remplacer *spin* par *moment cinétique* dans l'affirmation traduite par (20.110), et inversement.

[60] Une autre factorisation apparaît dans [15].

[61] i.e. la diagonale SW - NE.

[62] En dimension 2, cet opérateur n'est autre que σ_x.

[63] Dans un espace vectoriel, il existe une infinité de racines de l'unité. Par exemple, en dimension 2, on connaît déjà les 8 matrices : $\mathbf{1}$, les trois matrices de Pauli et leurs opposés. Cela étant, toutes les matrices s'en déduisant par une transformation (unitaire par exemple) sont encore des racines de l'unité.

Dès lors, l'opérateur K s'écrit :

$$\boxed{K = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\pi J_z}\,\sqrt{\mathbf{1}}} \tag{20.116}$$

Dans le cas particulier $S = 1/2$, $\sqrt{\mathbf{1}} = \sigma_x$ et on a simplement :

$$K = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\pi S_z}\,\sigma_x = \mathrm{e}^{-\frac{\mathrm{i}\pi}{2}\sigma_z}\,\sigma_x = -\mathrm{i}\sigma_z\sigma_x = \sigma_y \ , \tag{20.117}$$

soit explicitement $K|\pm\frac{1}{2}\rangle = \pm\mathrm{i}\,|\mp\frac{1}{2}\rangle$, en conformité avec (20.105). ♦

20.5 Équation de Dirac

20.5.1 Émergence du spin dans un cadre relativiste

Il s'agit ici de présenter, de façon élémentaire, les arguments qui, historiquement, ont attribué au spin une origine relativiste. Afin de rester au niveau le plus simple et pour aller droit au but, on considérera d'abord le cas d'une particule *libre*, un électron pour fixer les idées. Cette version première[64] de la théorie est due à Dirac [76] et s'appelle historiquement *Théorie de l'électron de Dirac*. Par la suite, on examinera le cas de l'électron lié au sein de l'atome d'hydrogène. Enfin, l'analyse de la limite faiblement relativiste permettra d'introduire le Hamiltonien de Pauli et fournira la justification précise de l'expression de l'interaction spin-orbite donnée en (20.6).

L'approche adoptée ici évite délibérément le formalisme covariant afin de rendre la lecture aussi accessible que possible. Bien évidemment, une fois obtenue l'équation de Dirac par des moyens pragmatiques et quelque peu heuristiques, il convient d'en vérifier l'invariance au sens d'une transformation de Lorentz ; ce point est abondamment discuté dans les traités classiques [15], [77], [78], auxquels est renvoyé le lecteur voulant en savoir (beaucoup) plus.

Par ailleurs, la suite montrera clairement comment cette théorie introduit spontanément un certain degré de liberté supplémentaire dont toutes les propriétés physiques autorisent qu'il soit identifié au spin de l'électron, et c'est pourquoi on évoque parfois la *nature relativiste du spin*. Si cette qualification semble historiquement légitime – et a prévalu quelque temps –, l'analyse au fond montre qu'elle n'est toutefois pas justifiée, pour des raisons à élucider mais qui sont incontestables : le spin est associé aux *spineurs*, objets introduits par Cartan (voir p. 856) sur des considérations exclusivement basées sur

[64] Il est intéressant de savoir que, chronologiquement, la première équation quantique relativiste est due à ... Schrödinger (section 6 du 4e article de 1926 [67]). Selon Dirac [72], Schrödinger était déçu par cette équation qui, au contraire du calcul effectué par Sommerfeld dans le cadre de l'Ancienne Théorie des Quanta (voir Tome I, sous-section 7.3.3), ne donnait pas la bonne structure fine pour la raie H_α, le *splitting* trouvé par Schrödinger étant presque trois fois trop grand. L'équation obtenue par Schrödinger est en fait celle qui fut redécouverte par Klein [73], Fock [74] et Gordon [75] très peu de temps après. C'est pourquoi, dans la littérature [53], l'équation dite de Klein - Gordon s'appelle parfois *équation relativiste de Schrödinger*.

l'analyse du groupe des rotations, n'ayant de ce fait aucun rapport avec des prémisses relativistes au sens d'Einstein[65].

Cela étant réalisé, il reste à comprendre pourquoi la théorie de Dirac, de fait, produit le spin ; l'argument est schématiquement le suivant. La théorie doit être invariante dans toute transformation de Lorentz, généralisation, quand la vitesse c est finie, de la transformation de Galilée. Or il se trouve que le produit de deux transformations de Lorentz n'est une transformation de Lorentz que si toutes les vitesses impliquées par les changements de repères sont parallèles entre elles. Autrement dit, pour avoir le groupe de symétrie *complet*, force est d'introduire explicitement et simultanément les *rotations* (voir par exemple [12], éq. III.55) ; cette implication obligatoire des rotations ne se produit pas pour les transformations de Galilée où la règle de composition des vitesses est la simple *somme vectorielle*. Cela étant fait, on incorpore implicitement toutes les nécessités liées à une telle symétrie, et en particulier toutes les formes de représentations, de dimension impaire *et* paire, respectivement associées aux valeurs entières *et* demi-entières du moment cinétique. En définitive, c'est l'implication nécessaire du groupe des rotations dans le cadre relativiste d'Einstein qui permet de *ramasser* le spin, inscrit sans doute aucun dans la seule symétrie de rotation.

On verra que la formalisation de Dirac soulève de grandes difficultés, et conduit finalement à une impasse dont on ne peut sortir qu'en faisant demi-tour et en abandonnant le cadre d'une théorie à *une* particule. Le premier symptôme grave, survenant dès le cas le plus simple (l'électron libre) est certainement la nécessité, dans une telle théorie, d'introduire en bout de course une... infinité de particules constituant la *mer de Dirac* (voir p. 888). Une autre difficulté apparaît également dans le traitement d'un atome hydrogénoïde de numéro atomique Z, lorsque le paramètre $Z\alpha$ devient de l'ordre de 1 (voir p. 897). De fait, dès que l'on pousse la théorie de Dirac, les difficultés se multiplient et les paradoxes apparaissent (par exemple celui de Klein). Leur accumulation, tout autant que les impossibilités à les lever sans voir surgir de nouvelles incohérences, ont finalement conduit à une reformulation complète de la Mécanique quantique relativiste, constituant la *Théorie quantique des champs* où, fondamentalement, le nombre de particules n'est pas conservé et, notamment, incorpore d'emblée la possibilité de création de paires particules – antiparticules. Qu'il faille, en un sens, *désapprendre* la théorie de Dirac pour pouvoir aborder la Théorie des champs n'est pas toutefois une raison suffisante pour renoncer à en donner un exposé élémentaire : son importance n'est pas seulement de nature historique, elle réside aussi dans le fait qu'il s'agit d'un nouvel exemple d'édification d'une théorie physique accomplie en réunissant avec une grande élégance quelques principes fondamentaux – l'extrême attention que portait Dirac à la *beauté* d'une théorie a déjà été mentionnée (voir Tome I, p. 388). De surcroît, la constatation implacable de l'impasse dans laquelle elle conduit, tout autant que la nécessité d'en sortir, ont permis des avancées majeures, dont l'importance est certainement comparable à celles dont nous sommes redevables à Heisenberg et Schrödinger, eux aussi confrontés à des difficultés insurmontables dans le cadre de pensée classique, à l'évidence et définitivement devenu archaïque.

[65] Rappelons que la notion de *relativité* est tout autant présente dans l'invariance au sens de Galilée.

20.5.2 Construction de l'équation de Dirac

A priori – et en feignant d'ignorer les difficultés à venir anticipées ci-dessus –, il est raisonnable de croire que la construction de l'équation d'onde relativiste peut être entreprise en suivant une procédure analogue à celle de son homologue non-relativiste. Tout compte fait, pour une particule libre non-relativiste, un racourci pour obtenir l'équation de Schrödinger en représentation-q :

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \Psi = \frac{1}{2m} (-\mathrm{i}\hbar\vec{\nabla})^2 \Psi \tag{20.118}$$

consiste à partir de l'égalité classique $H = \frac{\vec{p}^2}{2m}$ et à y faire la substitution formelle :

$$H \to \mathrm{i}\hbar \frac{\partial}{\partial t} , \qquad \vec{p} \to -\mathrm{i}\hbar \vec{\nabla} . \tag{20.119}$$

De façon analogue, pour le cas relativiste, on part du Hamiltonien classique (i.e. non quantique) pour une particule libre de masse[66] m :

$$H^2 = p^2 c^2 + m^2 c^4 ; \tag{20.120}$$

en y faisant la même substitution (20.119), on obtient :

$$\left(\mathrm{i}\hbar \frac{\partial}{\partial t}\right)^2 \Psi = \left[\left(- \mathrm{i}\hbar \vec{\nabla}\right)^2 c^2 + m^2 c^4\right] \Psi , \tag{20.121}$$

soit :

$$\boxed{\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta + \frac{m^2 c^2}{\hbar^2}\right) \Psi = 0} \tag{20.122}$$

La quantité $\frac{\hbar}{mc}$ apparaissant dans cette équation est la seule longueur disponible avec la constante de Planck, la masse de l'électron et la vitesse de la lumière ; c'est la longueur d'onde Compton de l'électron, qui vaut environ 0,024 Å (voir Tome I, section 5.3, éq. (5.44)). L'équation (20.122) s'appelle équation de Klein - Gordon [73], [75], obtenue également par Fock (et par... Schrödinger, voir note 64 p. 873). Elle conduit à des difficultés de deux ordres :

1. C'est une équation du *deuxième* ordre par rapport au temps, alors que jusqu'à présent on a admis que la connaissance de l'état à $t = 0$ suffit à déterminer l'état à tout instant ultérieur. La généralisation relativiste ne devrait pas exiger un changement de l'ordre de l'équation fondamentale, changement qui procède par tout ou rien (comme une symétrie, présente ou absente). On peut construire des développements en $1/c$ pour retrouver la limite non-relativiste : par continuité[67], on ne peut pas passer de l'ordre 2 à l'ordre 1.

[66] m désigne bien sûr la masse *propre* de la particule, parfois appelée *masse au repos* pour la distinguer de la masse $m\gamma$, où $\gamma \stackrel{\text{déf}}{=} \left[1 - (v/c)^2\right]^{-1/2}$.

[67] Bien sûr, la limite quantique non-relativiste pourrait être singulière, mais justement ce n'est pas le cas. Si le "petit paramètre" de la Mécanique quantique ($\hbar$!) apparaît dans le terme de plus haut degré dans l'équation de Schrödinger, faisant de la limite classique (non-quantique et non-relativiste) une limite singulière, rien de tel ne se produit en ce qui concerne la limite non-relativiste.

Toutefois, cette objection peut être contrée, au moins formellement. La réponse, même si elle n'est pas suffisante, est d'ailleurs éclairante et se révèlera féconde pour surmonter d'autres difficultés ultérieures. En effet, une équation du second ordre peut en fait résulter de la contraction de deux équations plus "fondamentales", toutes deux d'ordre 1. Un exemple trivial est fourni par un oscillateur (classique) à une dimension. L'équation dynamique est :

$$m\ddot{x} = -kx \ . \tag{20.123}$$

Elle est du second ordre, mais on sait bien qu'elle provient des deux équations de Hamilton :

$$\dot{x} = \frac{p}{m} \ , \qquad \dot{p} = -kx \ , \tag{20.124}$$

qui sont toutes deux du premier ordre ; on peut d'ailleurs écrire celles-ci sous forme matricielle :

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{bmatrix}\begin{bmatrix} x \\ p \end{bmatrix} \tag{20.125}$$

qui constitue bien un système du premier ordre, portant sur une description dynamique *complexifiée* (bien sûr, le nombre de conditions initiales n'a pas changé, il en faut toujours deux). De la même façon, on peut récrire l'équation de Klein - Gordon et mettre en évidence une fonction d'onde à *deux* composantes, en procédant comme suit. L'équation (20.122) s'écrit aussi :

$$\left(1 + \frac{\mathrm{i}\hbar}{mc^2}\frac{\partial}{\partial t}\right)\left(1 - \frac{\mathrm{i}\hbar}{mc^2}\frac{\partial}{\partial t}\right)\Psi = \frac{\hbar^2}{m^2c^2}\Delta\Psi \ , \tag{20.126}$$

forme qui suggère d'introduire les deux combinaisons :

$$\Psi_{\mathrm{a}} \stackrel{\text{déf}}{=} \Psi + \frac{\mathrm{i}\hbar}{mc^2}\frac{\partial\Psi}{\partial t} \ , \qquad \Psi_{\mathrm{b}} \stackrel{\text{déf}}{=} \Psi - \frac{\mathrm{i}\hbar}{mc^2}\frac{\partial\Psi}{\partial t} \ , \tag{20.127}$$

d'où inversement :

$$\Psi = \frac{1}{2}(\Psi_{\mathrm{a}} + \Psi_{\mathrm{b}}) \ , \qquad \frac{\partial\Psi}{\partial t} = \frac{mc^2}{2\mathrm{i}\hbar}(\Psi_{\mathrm{a}} - \Psi_{\mathrm{b}}) \ . \tag{20.128}$$

Dès lors, (20.126) peut s'écrire des deux façons :

$$\left(1 - \frac{\mathrm{i}\hbar}{mc^2}\frac{\partial}{\partial t}\right)\Psi_{\mathrm{a}} = \frac{\hbar^2}{2m^2c^2}\Delta(\Psi_{\mathrm{a}} + \Psi_{\mathrm{b}}) \ , \tag{20.129}$$

$$\left(1 + \frac{\mathrm{i}\hbar}{mc^2}\frac{\partial}{\partial t}\right)\Psi_{\mathrm{b}} = \frac{\hbar^2}{2m^2c^2}\Delta(\Psi_{\mathrm{a}} + \Psi_{\mathrm{b}}) \ , \tag{20.130}$$

soit, sous la forme matricielle :

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\begin{bmatrix} \Psi_{\mathrm{a}} \\ \Psi_{\mathrm{b}} \end{bmatrix} = \begin{bmatrix} mc^2 - \frac{\hbar^2}{2m}\Delta & -\frac{\hbar^2}{2m}\Delta \\ +\frac{\hbar^2}{2m}\Delta & -mc^2 + \frac{\hbar^2}{2m}\Delta \end{bmatrix}\begin{bmatrix} \Psi_{\mathrm{a}} \\ \Psi_{\mathrm{b}} \end{bmatrix} \ . \tag{20.131}$$

On retrouve ainsi formellement une équation d'évolution du premier ordre par rapport au temps – il reste toutefois que deux conditions initiales sont toujours à prescrire. En ce sens, la difficulté n'est résolue que *formellement*.

2. Il existe en réalité une autre difficulté, beaucoup plus grave si l'on s'en tient, comme en Mécanique quantique non-relativiste, à une théorie à une seule particule. L'équation (20.122) étant donnée, et si l'on veut maintenir l'interprétation usuelle de Ψ en tant qu'amplitude de probabilité (pour avoir une probabilité fabriquée avec $|\Psi|^2$), la cohérence exige de montrer qu'il existe une équation de conservation *découlant* de l'équation fondamentale, à savoir (20.122) – on se souvient que l'équation de Schödinger entraîne l'équation de conservation $\partial_t \rho + \mathrm{div}\vec{j} = 0$ (voir Tome I, p. 289). En gardant comme courant l'expression symétrisée que l'on peut construire avec Ψ, Ψ^* et la vitesse $\vec{v}$, soit $\frac{1}{m}\vec{p} \equiv \frac{\hbar}{\mathrm{i}m}\vec{\nabla}$, on pose à nouveau :

$$\vec{j} \overset{\text{déf}}{=} \frac{\hbar}{2\mathrm{i}m}\left(\Psi^*\vec{\nabla}\Psi - \Psi\vec{\nabla}\Psi^*\right) . \tag{20.132}$$

En ce qui concerne la densité ρ, que l'on attend quadratique en Ψ et Ψ^*, on constate que le simple produit $\Psi\Psi^*$ ne peut pas convenir pour reconstituer une équation de conservation puisque celle-ci contient une dérivée première en temps, et que l'équation de Klein - Gordon introduit une dérivée *seconde* : la densité doit donc être fabriquée à partir de Ψ *et* de $\partial_t\Psi$. Une démarche par essai et erreur montre que si l'on pose :

$$\rho \overset{\text{déf}}{=} \frac{\mathrm{i}\hbar}{2mc^2}\left(\Psi^*\frac{\partial\Psi}{\partial t} - \Psi\frac{\partial\Psi^*}{\partial t}\right) , \tag{20.133}$$

alors l'équation de Klein - Gordon (20.122) prend effectivement la forme souhaitée d'une équation de conservation :

$$\boxed{\frac{\partial\rho}{\partial t} + \mathrm{div}\,\vec{j} = 0} \tag{20.134}$$

Le seul ennui est que si la densité ρ définie en (20.133) est bien réelle (par construction)... elle n'a aucune raison d'être positive : l'équation étant du second ordre en temps, les deux facteurs $\mathrm{e}^{\pm(1/\mathrm{i}\hbar)Et}$ sont possibles, les solutions peuvent avoir une énergie positive ou négative ; la densité ρ définie en (20.133) est positive pour les solutions stationnaires[68] à énergie positive, négative pour les autres, et ne saurait donc représenter une densité de probabilité de présence pour *une* particule.

L'équation de Klein - Gordon ne peut être retenue qu'au prix d'une réinterprétation, qui a été initiée par Pauli et Weisskopf ([15], p. 764, [79], p. 76). Fondamentalement, en multipliant $\vec{j}$ et ρ par une charge, ceux-ci représentent alors un courant et une densité de charge. Comme, de toute façon, le nombre de particules ne saurait être fixé au sein d'une théorie relativiste (possibilité de création de paires), $\vec{j}$ et ρ peuvent être associés à la différence entre le nombre de charges positives et le nombre de charges négatives ([80], p. 195). La théorie devient alors une théorie à une charge (la charge totale), et perd son statut de théorie à une particule. L'équation de Klein - Gordon (20.122) reste, en Théorie quantique des champs, l'équation fondamentale pour un champ de bosons de spin nul [78]. Ses solutions réelles décrivent des

[68] Tout comme dans le cas non-relativiste, les états stationnaires de l'équation de Klein - Gordon sont de la forme $\mathrm{e}^{(1/\mathrm{i}\hbar)Et}\psi(\vec{r})$ où l'énergie E, précisément, peut être positive ou négative puisque l'équation est du second ordre en temps.

particules neutres, les solutions complexes décrivent des particules avec une charge électrique, et leurs antiparticules. La densité ρ est alors interprétable comme la différence entre les densités de particules et d'antiparticules.

De toute façon, l'équation de Klein - Gordon n'est pas satisfaisante pour l'électron, en dépit de la présence, dans sa version (20.131), des deux combinaisons $\Psi_{\rm a}$ et $\Psi_{\rm b}$ (qui pourraient faire penser aux deux composantes d'un spineur associé à $S = 1/2$). En effet, si l'on examine la limite non-relativiste ([80], p. 199 et ci-après p. 888), on trouve que l'une des composantes est d'ordre 0 en v/c, alors que l'autre est d'ordre 1 : elles sont donc loin de jouer un rôle symétrique, alors que, même dans une théorie pseudo-relativiste à la Pauli, les deux composantes du spineur doivent être du même ordre de grandeur. Au total, la limite non-relativiste produit finalement une fonction d'onde à *une* composante, et ne convient donc pas pour l'électron (au mieux, l'équation de Klein - Gordon convient *a priori* pour une particule de spin *nul*).

S'agissant de construire une équation relativiste, il faut avant tout s'ancrer sur l'obligation de faire jouer des rôles symétriques aux coordonnées d'espace et au temps. En particulier, si l'on maintient l'idée d'une équation d'ordre 1 en temps, les coordonnées doivent également figurer par des dérivées du premier ordre. Quand on revient à la forme "linéarisée" de Klein - Gordon (20.131), on devine ce qu'il faut tenter de faire. Dans la matrice au second membre, apparaît le Laplacien (dérivées spatiales du deuxième ordre) ; pour se ramener à des dérivées du premier ordre, on peut jouer au même jeu qu'avec le temps dans le cas de l'oscillateur harmonique – ce qui a permis de passer de (20.123) à (20.125) –, c'est-à-dire à *complexifier* à nouveau la représentation, mais en impliquant cette fois les variables spatiales. Ainsi survient la suggestion forte de doubler la dimension de la matrice 2×2, ce qui revient à dire que chaque combinaison $\Psi_{\rm a}$ et $\Psi_{\rm b}$ est en fait elle-même un *doublet*, composantes de l'élément d'un espace vectoriel à deux dimensions, produisant finalement une fonction d'onde à *quatre* composantes ; en retenant cette hypothèse de travail, celles-ci sont désormais notées Ψ_μ, $\mu = 1, 2, 3, 4$, que l'on peut écrire comme une matrice unicolonne :

$$\Psi = \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{bmatrix} . \qquad (20.135)$$

Dès lors, on posera que la densité de probabilité est $\rho(\vec{r}, t) = \sum_{\mu=1}^{4} |\Psi_\mu(\vec{r}, t)|^2$. Cela étant retenu, l'équation (du premier ordre en temps et espace) la plus simple que l'on peut former est une *forme linéaire* de $\mathrm{i}\hbar\frac{\partial}{\partial t}$ et de $\vec{p} = -\mathrm{i}\hbar\vec{\nabla}$; par ailleurs, pour une vitesse nulle, l'énergie doit se réduire à un terme contenant l'énergie dite de repos mc^2. Au total, on peut *a priori* poser la forme suivante[69] :

$$\boxed{\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(\vec{r}, t) = (c\vec{\alpha}.\vec{p} + \beta mc^2)\Psi(\vec{r}, t) \equiv H_{\rm D}\Psi(\vec{r}, t)} \qquad (20.136)$$

[69]La notation $\vec{\alpha}$ est traditionnelle : ne pas confondre avec la constante de structure fine $\alpha \simeq \frac{1}{137}$, pourtant incontournable dans le présent contexte...

H_D est par définition le Hamiltonien de Dirac ; $\vec{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$ est un ensemble de trois objets indépendants du temps et de l'espace, agissant sur des vecteurs à quatre composantes : ce sont des matrices[70] constantes 4×4, tout comme β, qui doivent manifestement être hermitiques. $\vec{\alpha}$ et β sont sans dimension physique ; ces matrices mélangent entre elles les quatre composantes Ψ_μ. En présence d'un champ de forces statique et scalaire associé à l'énergie potentielle $V(\vec{r})$, la généralisation de (20.136) est[71] :

$$\boxed{i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},\, t) = [c\vec{\alpha}.\vec{p} + V(\vec{r})\mathbf{1}_4 + \beta mc^2]\Psi(\vec{r},\, t)} \tag{20.137}$$

En divisant membre à membre par $i\hbar$, l'équation (20.136) s'écrit :

$$\Big(\frac{1}{c}\frac{\partial}{\partial t} + \vec{\alpha}.\vec{\nabla} + \frac{imc}{\hbar}\beta\Big)\Psi(\vec{r},\, t) = 0 \ , \tag{20.138}$$

soit, sous forme plus explicite :

$$\Big(\frac{1}{c}\frac{\partial}{\partial t} + \sum_{u=x,\,y,\,z} \alpha_u\frac{\partial}{\partial u} + \frac{imc}{\hbar}\beta\Big)\Psi(\vec{r},\, t) = \mathbf{0} \ . \tag{20.139}$$

Toute la question est maintenant de trouver les matrices $\vec{\alpha}$ et β. Il convient bien sûr de maintenir le lien avec l'expression classique (20.120), qui est vraie pour *toute* particule et qui produit l'équation de Klein - Gordon (20.122) par simple substitution des opérateurs aux grandeurs classiques[72]. Une façon de maintenir le contact est de faire des manipulations sur l'opérateur entre crochets dans (20.139) pour lui donner l'allure de l'opérateur apparaissant dans (20.122). Au total, il s'agit de trouver une marche à suivre permettant d'écrire des relations entre les matrices $\vec{\alpha}$ et β cherchées et, finalement, de les déterminer.

Pour exhiber à partir de (20.139) un premier membre ressemblant à l'équation de Klein - Gordon, il faut fabriquer des dérivées secondes en temps et en espace (images opératorielles de la relation classique (20.120), qui est quadratique en H et $\vec{p}$), sans pour autant faire apparaître de dérivées croisées temps - espace ; compte tenu de la forme $A + B$ de l'opérateur entre crochets au premier membre de (20.139), un peu de réflexion montre qu'il faut introduire l'opérateur de la forme $A - B$, soit :

$$\frac{1}{c}\frac{\partial}{\partial t} - \sum_{v=x,\,y,\,z} \alpha_v\frac{\partial}{\partial v} - \frac{imc}{\hbar}\beta \ , \tag{20.140}$$

[70] Il est d'usage de parler des *matrices de Dirac* car, le plus souvent, il est fait référence implicitement à la représentation standard. En réalité, $\vec{\alpha}$ et β sont des *opérateurs*, uniquement et entièrement caractérisés par leurs relations algébriques de définition, (20.143) et (20.144).

[71] Afin de simplifier les notations, et quand le contexte ne laissera place à aucune ambiguïté, $\mathbf{1}$ notera la matrice identité, qu'elle soit 2×2, 4×4...

[72] Pour cette raison, l'équation de Klein - Gordon apparaît comme un point de passage obligé entre la relation classique (i.e. non quantique) et l'équation cherchée.

et le faire agir sur le premier membre de (20.139) ; procédant ainsi, il vient :

$$\Big(\frac{1}{c}\frac{\partial}{\partial t} - \sum_v \alpha_v \frac{\partial}{\partial v} - \frac{imc}{\hbar}\beta\Big)\Big(\frac{1}{c}\frac{\partial}{\partial t} + \sum_u \alpha_u \frac{\partial}{\partial u} + \frac{imc}{\hbar}\beta\Big)\Psi(\vec{r},\, t) = 0 \ . \qquad (20.141)$$

Il suffit maintenant d'expliciter les termes au premier membre, et de jouer avec les indices muets de sommation pour obtenir[73] :

$$\Big[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{1}{2}\sum_{u,v}(\alpha_u\alpha_v + \alpha_v\alpha_u)\frac{\partial^2}{\partial u \partial v} - \frac{imc}{\hbar}\sum_u(\alpha_u\beta + \beta\alpha_u)\frac{\partial}{\partial u} + \big(\frac{mc}{\hbar}\big)^2\beta^2\Big]\Psi(\vec{r},\, t) = 0 \ ; \qquad (20.142)$$

alors cette équation se réduit à celle de Klein - Gordon ssi $\alpha_u\alpha_v + \alpha_v\alpha_u = 2\delta_{uv}$, avec $u,\, v = x,\, y,\, z$, soit :

$$\boxed{\alpha_u^2 = \mathbf{1}_4 \quad \forall u \qquad \alpha_u\alpha_v + \alpha_v\alpha_u = 0 \quad (u \neq v)} \qquad (20.143)$$

et :

$$\boxed{\beta^2 = \mathbf{1}_4 \qquad \alpha_u\beta + \beta\alpha_u = 0 \quad \forall u} \qquad (20.144)$$

Les α_u et β sont bien des *opérateurs* : les relations ci-dessus ne sauraient être satisfaites par des nombres ordinaires[74].

Les relations (20.143) et (20.144) étant admises, il n'est pas difficile de montrer qu'il existe une représentation où $\vec{\alpha}$ et β sont représentés par les matrices hermitiques 4×4 suivantes, écrites sous la forme[75] de blocs 2×2 :

$$\boxed{\vec{\alpha} = \begin{bmatrix} \mathbf{0} & \vec{\sigma} \\ \vec{\sigma} & \mathbf{0} \end{bmatrix} \qquad \beta = \begin{bmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_2 \end{bmatrix}} \qquad (20.145)$$

où les $\vec{\sigma}$ sont les matrices de Pauli et où $\mathbf{1}_2$ est la matrice identité 2×2. L'ensemble des quatre matrices données par (20.145) est appelé *représentation-standard*, étant entendu qu'il existe une infinité de couples $(\vec{\alpha},\, \beta)$ vérifiant (20.143) et (20.144), déductibles de (20.145) par une transformation unitaire quelconque ; comme toujours, ce sont les relations algébriques entre *opérateurs* qu'il convient de respecter, le choix d'une représentation étant une pure affaire de convenance dans une situation donnée. En tout cas, dans cette représentation, le Hamiltonien de Dirac H_{D}, pour une particule libre, a l'expression matricielle[76] 4×4 :

$$\boxed{H_{\mathrm{D}} = \begin{bmatrix} mc^2\mathbf{1}_2 & c\vec{\sigma}.\vec{p} \\ c\vec{\sigma}.\vec{p} & -mc^2\mathbf{1}_2 \end{bmatrix} \equiv mc^2\beta + c\vec{\alpha}.\vec{p}} \qquad (20.146)$$

[73] *A priori* les matrices α_u ne commutent pas entre elles, alors que $\frac{\partial^2}{\partial u \partial v} = \frac{\partial^2}{\partial v \partial u}$.

[74] On connaît cependant des nombres qui obéissent à ce type d'algèbre, ou qui, en tout cas anticommutent : ce sont les nombres (variables) de Grassmann, souvent notés ξ_l et tels que $\xi_l\xi_{l'} + \xi_{l'}\xi_l = 0$. Ces nombres, d'apparence exotique, apparaissent dans la formulation en intégrale fonctionnelle (*à la* Feynman) des systèmes de fermions.

[75] $\mathbf{0}$ note ici un bloc matriciel 2×2 identiquement nul.

[76] Pour la simplicité des notations, on continue à utiliser le même symbole pour un opérateur et pour sa matrice sur une base donnée.

Notons que le carré de H_D n'est autre que la matrice $(m^2c^4 + \vec{p}^2c^2)\mathbf{1}_4$, en résultat du fait que l'on a imposé aux quatre composantes Ψ_μ de satisfaire (20.122). Pour une particule d'énergie potentielle $V(\vec{r})$, on a :

$$\boxed{H_\mathrm{D} = \begin{bmatrix} [mc^2 + V(\vec{r})]\mathbf{1}_2 & c\vec{\sigma}.\vec{p} \\ c\vec{\sigma}.\vec{p} & [-mc^2 + V(\vec{r})]\mathbf{1}_2 \end{bmatrix} \equiv V(r)\mathbf{1}_4 + mc^2\beta + c\vec{\alpha}.\vec{p}} \tag{20.147}$$

Avant d'exploiter ces équations et de commencer les calculs, une brève analyse qualitative est utile, considérant le cas le plus simple d'une particule libre, (20.146). La structure de la matrice de H_D montre que dans la limite $\vec{p} \to 0$, les quatre composantes se découplent, deux d'entre elles ayant l'énergie $+mc^2$, les autres ayant l'énergie $-mc^2$; ce dernier résultat est surprenant, car on ne voit pas bien ce que peut représenter une telle solution à énergie *négative* : il s'agit là du premier symptôme de la difficulté foncière de la théorie de Dirac. Quand $\vec{p} \neq 0$, mais si $\|\vec{p}\|c \ll mc^2$, il n'y a pas de découplage au sens strict, mais on s'attend à ce que les énergies propres soit à nouveau voisines[77] de $\pm mc^2$; la fonction d'onde a cette fois quatre composantes non nulles, mais, visiblement, pour $E \sim +mc^2$, les deux premières seront grandes, les deux dernières seront petites ; l'inverse se produit pour l'autre solution $E \sim -mc^2$. C'est pourquoi, comme on le verra dans la suite, on parle respectivement dans chaque cas de *grandes* et de *petites* composantes.

Montrons tout de suite que l'équation de Dirac contient effectivement l'équation de conservation *nécessaire* $\partial_t\rho + \mathrm{div}\vec{j} = 0$. En introduisant les quatre composantes $\Psi_\mu(\vec{r}, t)$, la densité est $\rho(\vec{r}, t) = \langle\Psi(\vec{r}, t)|\Psi(\vec{r}, t)\rangle = \sum_\mu \Psi_\mu^*(\vec{r}, t)\Psi_\mu(\vec{r}, t)$. En multipliant par Ψ_μ^* chacune des quatre équations $\mathrm{i}\hbar\partial_t\Psi_\mu = \sum_{\mu'} H_{\mathrm{D}\,\mu\mu'}\Psi_{\mu'}$ et en en faisant la somme membre à membre, on obtient :

$$\mathrm{i}\hbar\sum_\mu \Psi_\mu^*\frac{\partial\Psi_\mu}{\partial t} = \sum_{\mu\mu'} \Psi_\mu^*(mc^2\beta + c\vec{\alpha}.\vec{p} + V(\vec{r})\mathbf{1}_4)_{\mu\mu'}\Psi_{\mu'} \ . \tag{20.148}$$

On prend maintenant la complexe conjuguée de cette équation, et on fait la différence membre à membre des deux ; les termes en $mc^2\beta + V(\vec{r})\mathbf{1}_4$ se compensent et il reste ($\vec{p} = -\mathrm{i}\hbar\vec{\nabla}$) :

$$\mathrm{i}\hbar\frac{\partial\rho}{\partial t} = -\mathrm{i}\hbar c\sum_{\mu\mu'}\left[\Psi_\mu^*\vec{\alpha}_{\mu\mu'}.\vec{\nabla}\Psi_{\mu'} + (\Psi_\mu^*\vec{\alpha}_{\mu\mu'}.\vec{\nabla}\Psi_{\mu'})^*\right] \ . \tag{20.149}$$

La quantité complexe conjuguée du crochet de droite est :

$$(\Psi_\mu^*\vec{\alpha}_{\mu\mu'}.\vec{\nabla}\Psi_{\mu'})^* = \Psi_\mu\vec{\alpha}_{\mu\mu'}^*.\vec{\nabla}\Psi_{\mu'}^* = \Psi_\mu\vec{\alpha}_{\mu'\mu}\vec{\nabla}\Psi_{\mu'}^* \ . \tag{20.150}$$

[77] Il s'agit d'un argument de type perturbatif, que l'on retrouvera systématiquement dans le chapitre 23. L'idée est que quand les éléments non-diagonaux sont petits devant la *différence* d'énergie entre états non-perturbés, les corrections d'énergie restent petites à condition que la perturbation soit non-singulière. Ici, cela signifie que $\|\vec{p}\|c \ll 2mc^2$, soit $\|\vec{p}\| \ll mc$: il s'agit clairement de la limite faiblement relativiste.

Une fois cette dernière expression réintroduite dans la sommation symétrique $\sum_{\mu\mu'}$, on peut échanger les deux indices muets, de sorte que (20.149) devient :

$$\frac{\partial\rho}{\partial t} = -c\sum_{\mu\mu'}\left(\Psi_\mu^*\vec{\alpha}_{\mu\mu'}.\vec{\nabla}\Psi_{\mu'} + \Psi'_\mu\vec{\alpha}_{\mu\mu'}\vec{\nabla}\Psi_\mu^*\right) \equiv -c\sum_{\mu\mu'}\vec{\alpha}_{\mu\mu'}.\left(\Psi_\mu^*\vec{\nabla}\Psi_{\mu'} + \Psi'_\mu\vec{\nabla}\Psi_\mu^*\right) \ . \tag{20.151}$$

La parenthèse à droite est $\vec{\nabla}(\Psi_\mu^*\Psi_{\mu'})$, d'où :

$$\frac{\partial\rho}{\partial t} = -c\vec{\nabla}\sum_{\mu\mu'}\Psi_\mu^*(\vec{r},\, t)\vec{\alpha}_{\mu\mu'}\Psi_{\mu'}(\vec{r},\, t) \equiv -c\vec{\nabla}\langle\Psi(\vec{r},\, t)|\vec{\alpha}|\Psi(\vec{r},\, t)\rangle \ . \tag{20.152}$$

L'équation de Dirac conduit bien à l'équation de conservation, densité et courant ayant les expressions :

$$\boxed{\rho(\vec{r},\, t) = \langle\Psi(\vec{r},\, t)|\Psi(\vec{r},\, t)\rangle \qquad \vec{j}(\vec{r},\, t) = c\langle\Psi(\vec{r},\, t)|\vec{\alpha}|\Psi(\vec{r},\, t)\rangle} \tag{20.153}$$

Comme il se doit, ces expressions redonnent leurs équivalents non-relativistes dans la bonne limite (voir problème 20.8.4). La forme du courant s'éclaire bien par référence à la vitesse, telle qu'elle est révélée par l'équation de Heisenberg pour $\vec{r}$; pour sa composante u ($u = x,\, y,\, z$), on a :

$$\mathrm{i}\hbar\frac{\mathrm{d}u_\mathrm{H}}{\mathrm{d}t} = [u,\, mc^2\beta + V(\vec{r})\mathbf{1}_4 + c\vec{\alpha}.\vec{p}]_\mathrm{H} = \mathrm{i}\hbar c\alpha_{\mathrm{H}\,u}(t) \ , \tag{20.154}$$

d'où :

$$\boxed{\vec{v}(t) \stackrel{\text{déf}}{=} \frac{\mathrm{d}\vec{r}_\mathrm{H}}{\mathrm{d}t} = c\,\vec{\alpha}_\mathrm{H}(t)} \tag{20.155}$$

(voir aussi problème 20.8.9). En conclusion, la théorie de Dirac d'une part introduit une densité positive, d'autre part contient en soi l'équation de conservation assurant, sur ce plan, la cohérence obligatoire.

Ce point nécessaire étant acquis, on peut maintenant établir le premier résultat spectaculaire de la théorie de Dirac, à savoir l'apparition *spontanée* d'une observable qui va pouvoir être identifiée avec le spin de l'électron ; la suite montre en effet que l'opérateur :

$$\vec{\Sigma} \stackrel{\text{déf}}{=} \begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix} \tag{20.156}$$

a toutes les qualités requises pour assurer que le vecteur :

$$\boxed{\vec{J} \stackrel{\text{déf}}{=} \vec{L} + \frac{\hbar}{2}\vec{\Sigma}} \tag{20.157}$$

qui se veut être un moment cinétique puisqu'il incorpore déjà le moment cinétique orbital, est effectivement une constante du mouvement lorsque le Hamiltonien est celui qui figure au second membre de (20.137), avec un champ sphérique $V(\vec{r}) \equiv V(r)$. Le terme potentiel $V(r)\mathbf{1}_4$ commute avec les matrices $\vec{\alpha}$ et β ; de surcroît, étant à symétrie

sphérique, il commute avec $\vec{L}$. Il en résulte que le commutateur $[\vec{J},\, H_{\mathrm{D}}]$ se réduit à $[\vec{J},\, H_{\mathrm{D}}] = [\vec{J},\, c\vec{\alpha}.\vec{p} + \beta mc^2]$, exactement comme pour une particule *libre*.

Calculons d'abord le commutateur de $\vec{L}$ avec la partie cinétique de H_{D}. Dans l'espace à 4 dimensions, le moment cinétique orbital est simplement :

$$\vec{L} = \begin{bmatrix} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_2 \end{bmatrix} (\vec{r} \times \vec{p}) \equiv (\vec{r} \times \vec{p})\mathbf{1}_4 \ , \tag{20.158}$$

puisque $\vec{p}$ agit de la même façon sur chaque composante ($\vec{p}\,\Psi_\mu = -\mathrm{i}\hbar\vec{\nabla}\Psi_\mu$). Il en résulte :

$$[\vec{L},\, H_{\mathrm{D}}] = [(\vec{r} \times \vec{p})\mathbf{1}_4,\, c\vec{\alpha}.\vec{p} + \beta mc^2] = [(\vec{r} \times \vec{p})\mathbf{1}_4,\, c\vec{\alpha}.\vec{p}] \ . \tag{20.159}$$

Par exemple, on a :

$$\big[L_z, H_{\mathrm{D}}\big] = c\Big[xp_y - yp_x, \sum_u \alpha_u p_u\Big] = \mathrm{i}\hbar c\,(\alpha_x p_y - \alpha_y p_x) \equiv \mathrm{i}\hbar c\,(\vec{\alpha} \times \vec{p})_z \ , \tag{20.160}$$

et de même pour les autres composantes, d'où :

$$\boxed{\big[\vec{L},\, H_{\mathrm{D}}\big] = \mathrm{i}\hbar c\,(\vec{\alpha} \times \vec{p})} \tag{20.161}$$

Physiquement, ceci signifie que $\vec{L}$ n'est pas une constante du mouvement[78], laissant pressentir qu'une interaction du genre spin-orbite est déjà contenue dans le formalisme – quoique non encore apparente (elle ressortira lors de la séparation de la variable radiale, voir (20.204)). Par ailleurs[79] :

$$\big[\vec{\Sigma},\, H_{\mathrm{D}}\big] = \left[\begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix}, \begin{bmatrix} mc^2\mathbf{1}_2 & c\vec{\sigma}.\vec{p} \\ c\vec{\sigma}.\vec{p} & -mc^2\mathbf{1}_2 \end{bmatrix}\right] = \begin{bmatrix} \mathbf{0} & c[\vec{\sigma},\, \vec{\sigma}.\vec{p}] \\ c[\vec{\sigma},\, \vec{\sigma}.\vec{p}] & \mathbf{0} \end{bmatrix} . \tag{20.162}$$

On a, par exemple :

$$[\sigma_x, \vec{\sigma}.\vec{p}] = [\sigma_x, \sum_u \sigma_u p_u] = 2\mathrm{i}(p_y\sigma_z - p_z\sigma_y) \equiv 2\mathrm{i}(\vec{p} \times \vec{\sigma})_x \ , \tag{20.163}$$

d'où :

$$\big[\vec{\Sigma},\, H_{\mathrm{D}}\big] = 2\mathrm{i}c\begin{bmatrix} \mathbf{0} & \vec{p} \times \vec{\sigma} \\ \vec{p} \times \vec{\sigma} & \mathbf{0} \end{bmatrix} \equiv 2\mathrm{i}c\,\vec{p} \times \vec{\alpha} \ . \tag{20.164}$$

À nouveau, ceci montre que $\vec{S}$ n'est pas une constante du mouvement. Prenant maintenant en compte (20.161) et (20.164), le commutateur de $\vec{J}$ défini en (20.157) avec le Hamiltonien de Dirac H_{D} est :

$$[\vec{J},\, H_{\mathrm{D}}] \equiv [\vec{L} + \frac{\hbar}{2}\vec{\Sigma},\, H_{\mathrm{D}}] = \mathrm{i}\hbar c(\vec{\alpha} \times \vec{p} + \vec{p} \times \vec{\alpha}) = 0 \ . \tag{20.165}$$

Ce résultat permet bien d'interpréter $(\hbar/2)\vec{\Sigma}$ comme le moment cinétique intrinsèque de l'électron : ajouté vectoriellement au moment cinétique orbital $\vec{L}$ pour former le moment

[78] Qu'en est-il de $\vec{L}^{\,2}$?

[79] Penser à utiliser la multiplication par bloc des matrices.

cinétique total $\vec{J}$, celui-ci est bien une constante du mouvement en présence d'un champ central. La théorie de Dirac engendre ainsi le spin sous la forme :

$$\boxed{\vec{S} = \frac{\hbar}{2}\vec{\Sigma} = \frac{\hbar}{2}\begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix}} \tag{20.166}$$

dans la représentation standard, le moment cinétique total étant $\vec{J} = \vec{L} + \vec{S}$.

Chronologiquement la première à faire émerger le spin sans ingrédient *ad hoc*, la théorie de Dirac a autorisé à croire que le spin lui-même est d'origine relativiste. En réalité, comme l'a montré Wigner, c'est plus précisément l'invariance de Lorentz de l'équation de Dirac – et l'inclusion forcée des rotations, voir p. 873 –, qui est à l'œuvre en coulisse. Quoi qu'il en soit, et comme annoncé plus haut, la théorie de Dirac fournit la valeur $g_e = 2$, comme on le voit en prenant la limite non-relativiste pour une particule libre en présence d'un champ électromagnétique. Cette limite ([15], p. 807), examinée en détail dans la sous-section 20.7 et où $\vec{S} = \hbar\vec{\sigma}/2$, fournit notamment le Hamiltonien :

$$H_{c\to\infty} = \frac{1}{2m}(\vec{p} - e\vec{A})^2 - \frac{e\hbar}{2m}\vec{\sigma}.\vec{\mathcal{B}} + eU \ . \tag{20.167}$$

appelé *Hamiltonien de Pauli*. Le second terme est de toute évidence le couplage magnétique avec le champ ; étant de la forme $-\frac{e\hbar}{2m}\vec{\sigma}.\vec{\mathcal{B}}$, il fait ressortir le moment magnétique propre de l'électron $\vec{\mu}_S = \frac{e}{m}\frac{\hbar}{2}\vec{\sigma}$, soit $\vec{\mu}_S = \frac{e}{m}\vec{S}$. Par comparaison avec (20.18) :

$$\boxed{\vec{\mu}_S = \frac{e}{m}\vec{S} = g_e\frac{e}{2m}\vec{S} \quad \Longleftrightarrow \quad g_e = 2} \tag{20.168}$$

Comme déjà mentionné (voir note 15 p. 853), ceci n'est pas strictement correct ; l'écart $g_e - 2 = 0,002\,319\,304\,3\ldots$ résulte de corrections *radiatives* et ne peut apparaître que dans une théorie traitant matière et photons sur un pied d'égalité, les englobant dans un même cadre quantique (c'est l'Électrodynamique quantique). Quoi qu'il en soit, la présence de $\vec{\sigma}$ dans (20.167) montre que le Hamiltonien de Pauli agit sur une fonction d'onde à *deux composantes* ; ainsi, la limite faiblement relativiste reproduit exactement le Hamiltonien Zeeman (20.60) – à l'écart $g_e - 2$ près, évidemment.

20.5.3 États stationnaires d'un électron libre

Pour un électron libre, il est aisé de trouver les états stationnaires en théorie de Dirac ; en posant $|\Psi(\vec{r}, t)\rangle = \mathrm{e}^{\frac{1}{i\hbar}Et}|\psi(\vec{r})\rangle$, (20.136) devient :

$$\boxed{E|\psi(\vec{r})\rangle = H_D|\psi(\vec{r})\rangle} \tag{20.169}$$

et constitue l'équation aux valeurs et fonctions propres de Dirac pour un électron libre. En l'absence de champ de forces, l'impulsion est une constante du mouvement ($\vec{p}$ commute

avec H_{D}), et on peut chercher des solutions sous la forme d'ondes planes, correspondant à une impulsion déterminée, c'est-à-dire fonctions propres de $\vec{p} = -\text{i}\hbar\vec{\nabla}$:

$$|\psi(\vec{r}\,)\rangle = \text{e}^{\text{i}\vec{k}.\vec{r}}\,|C(\vec{k}\,)\rangle \stackrel{\text{déf}}{=} |\psi^{(\vec{k})}(\vec{r}\,)\rangle \ . \tag{20.170}$$

Dit autrement, l'invariance par translation résultant de l'absence d'une force assure que les générateurs des translations (les trois composantes de $\vec{p}$) sont des constantes du mouvement. Dans (20.170), $|\psi^{(\vec{k})}(\vec{r}\,)\rangle$ et $|C(\vec{k}\,)\rangle$ sont des matrices-colonne à quatre éléments ; sous forme explicite :

$$\psi_\mu^{(\vec{k})}(\vec{r}\,) = \text{e}^{\text{i}\vec{k}.\vec{r}}\,C_\mu(\vec{k}\,) \qquad (\mu = 1,\, 2,\, 3,\, 4) \ , \tag{20.171}$$

où chaque constante C_μ, paramétrée par $\vec{k}$, est bien évidemment indépendante de $\vec{r}$. En reportant dans (20.169) et en utilisant $\vec{p}\,\text{e}^{\text{i}\vec{k}.\vec{r}} = -\text{i}\hbar\vec{\nabla}\text{e}^{\text{i}\vec{k}.\vec{r}} = \hbar\vec{k}\,\text{e}^{\text{i}\vec{k}.\vec{r}}$ ($\hbar\vec{k}$ est donc la valeur propre de $\vec{p}$), on obtient $(c\hbar\vec{k}.\vec{\alpha} + \beta mc^2)|C(\vec{k}\,)\rangle = E|C(\vec{k}\,)\rangle$, soit, écrivant par blocs 2×2 la matrice 4×4 de H_{D} :

$$\begin{bmatrix} mc^2\mathbf{1}_2 & c\hbar\vec{k}.\vec{\sigma} \\ c\hbar\vec{k}.\vec{\sigma} & -mc^2\mathbf{1}_2 \end{bmatrix} \begin{bmatrix} C_1(\vec{k}\,) \\ C_2(\vec{k}\,) \\ C_3(\vec{k}\,) \\ C_4(\vec{k}\,) \end{bmatrix} = E \begin{bmatrix} C_1(\vec{k}\,) \\ C_2(\vec{k}\,) \\ C_3(\vec{k}\,) \\ C_4(\vec{k}\,) \end{bmatrix} \iff H_{\text{D}}|C(\vec{k}\,)\rangle = E|C(\vec{k}\,)\rangle \, . \tag{20.172}$$

Pour trouver commodément l'énergie E, on commence par appliquer deux fois H_{D} à C :

$$H_{\text{D}}^2\,|C(\vec{k}\,)\rangle = E^2\,|C(\vec{k}\,)\rangle \ ; \tag{20.173}$$

en calculant (toujours par blocs) le carré de la matrice figurant dans (20.172), on trouve facilement la forme matricielle[80] de (20.173) :

$$\begin{bmatrix} (mc^2)^2\mathbf{1}_2 + \hbar^2c^2(\vec{k}.\vec{\sigma})(\vec{k}.\vec{\sigma}) & \mathbf{0} \\ \mathbf{0} & (mc^2)^2\mathbf{1}_2 + \hbar^2c^2(\vec{k}.\vec{\sigma})(\vec{k}.\vec{\sigma}) \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = E^2 \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} . \tag{20.174}$$

Compte tenu de (18.143), $(\vec{k}.\vec{\sigma})(\vec{k}.\vec{\sigma})$ est égal à $\vec{k}^2\mathbf{1}_2$; le carré de l'énergie est donc $E^2 = (mc^2)^2 + \hbar^2\vec{k}^2c^2$, comme il se doit puisque $\vec{p} = \hbar\vec{k}$: ceci signifie que l'énergie est donnée (au signe près) dès que $\hbar\vec{k}$ est fixé, autre conséquence du fait que, la particule étant libre, son impulsion $\vec{p} = \hbar\vec{k}$ est une constante du mouvement. On retrouve alors la relation $E^2 = (mc^2)^2 + (pc)^2$, p désignant le module de l'impulsion, d'où les deux formes possibles de la relation de dispersion :

$$E = \pm\sqrt{(mc^2)^2 + (pc)^2} \equiv \pm E_p \ , \tag{20.175}$$

E_p désignant la racine carrée positive. Ainsi, deux classes de solutions apparaissent, les unes d'énergie positive, supérieure ou égale à $E_{p=0} = mc^2$, les autres d'énergie négative,

[80]Dans la suite, la dépendance en $\vec{k}$ des C_μ est sous-entendue pour alléger les notations.

bornées supérieurement par $-E_{p=0} = -mc^2$. Ces deux classes sont séparées[81] l'une de l'autre par un *gap* égal à $2mc^2$, un fait dont l'interprétation due à Dirac sera donnée plus loin (voir p. 888).

Il s'avère commode de paramétriser comme suit :

$$p = mc\sinh\phi \ , \qquad E = \pm mc^2\cosh\phi \ , \qquad \phi \geq 0 \ , \tag{20.176}$$

et alors la limite non relativiste $p \ll mc \iff pc \ll mc^2$ correspond à $\phi \ll 1$. Par ailleurs, comme l'espace est de dimension 4, chaque valeur de l'énergie (à $\vec{k}$ donné et pour une classe) est dégénérée deux fois. Il existe donc une autre observable qui, jointe à $H_{\rm D}$, constitue un ECOC ; ses vecteurs propres seront trouvés en formant les bonnes combinaisons linéaires d'états de *même* énergie.

On peut toujours choisir l'axe Oz le long du vecteur $\hbar\vec{k}$ fixé, dont le module est p. Dès lors, (20.172) prend la forme explicite plus simple :

$$\begin{bmatrix} mc^2\mathbf{1}_2 & pc\sigma_z \\ pc\sigma_z & -mc^2\mathbf{1}_2 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = E \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} \ . \tag{20.177}$$

Chaque bloc (diagonal ou non-diagonal) étant lui-même une matrice diagonale, ce système 4×4 se dissocie en deux systèmes 2×2 découplés :

$$\begin{cases} (mc^2 - E)C_1 + pc\,C_3 = 0 \\ pc\,C_1 - (mc^2 + E)C_3 = 0 \end{cases} \ , \qquad \begin{cases} (mc^2 - E)C_2 - pc\,C_4 = 0 \\ -pc, C_2 - (mc^2 + E)C_4 = 0 \end{cases} \ . \tag{20.178}$$

Ces deux systèmes ont la *même* équation caractéristique en E, dont les racines sont données par les expressions (20.175) (cette identité n'est autre que la dégénérescence annoncée plus haut). Les valeurs propres étant déterminées, il reste à trouver les vecteurs propres. De toute évidence, $\vec{\Sigma}.\frac{\vec{p}}{p} = \Sigma_z$ commute avec $H_{\rm D}$, puisque (voir 20.162) :

$$[\Sigma_z, H_{\rm D}] = \left[\begin{bmatrix} \sigma_z & \mathbf{0} \\ \mathbf{0} & \sigma_z \end{bmatrix}, \begin{bmatrix} mc^2\mathbf{1}_2 & cp\sigma_z \\ cp\sigma_z & -mc^2\mathbf{1}_2 \end{bmatrix}\right] =$$
$$\begin{bmatrix} \mathbf{0} & c[\sigma_z, p\sigma_z] \\ c[\sigma_z, p\sigma_z] & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \ . \tag{20.179}$$

Σ_z, observable qui participe à l'ECOC, représente le cosinus de l'angle entre le spin $\vec{S}$ et la direction de propagation ; l'équation propre pour Σ_z s'écrit :

$$\begin{bmatrix} \sigma_z & \mathbf{0} \\ \mathbf{0} & \sigma_z \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = \Sigma \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} \qquad (\sigma = \pm 1) \ . \tag{20.180}$$

Cette équation montre que des combinaisons linéaires arbitraires des composantes 1 et 3 ou des composantes 2 et 4, sont respectivement associées aux valeurs propres $\Sigma = +1$

[81] Pour une particule de masse nulle (le photon par exemple), on a $E = \pm\hbar kc$ et le *gap* est nul.

et $\Sigma = -1$ de Σ_z. Les combinaisons linéaires qui sont *aussi* propres de $H_{\rm D}$ s'obtiennent en diagonalisant les deux systèmes (20.178), celui de gauche correspondant à $\Sigma = +1$, celui de droite à $\Sigma = -1$. Au total, désignant par ε le signe de l'énergie ($E = \varepsilon E_p$, voir (20.175)), les vecteurs propres auront pour étiquettes, à impulsion $\hbar\vec{k}$ donnée, le couple de nombres quantiques (ε, Σ) et notés $|C_{\varepsilon,\,\Sigma}\rangle$.

Pour les solutions à énergie positive, un calcul facile fournit les expressions des 4 composantes des $|C_{\varepsilon,\,\Sigma}\rangle$, soit finalement les vecteurs propres $\psi_{+,\,\Sigma}(\vec{r})$ (voir éq. 20.170) :

$$|\psi^{(\vec{k})}_{+,\,+1}(\vec{r})\rangle = \frac{\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}}{\sqrt{\cosh\phi}}\begin{bmatrix}\cosh\frac{\phi}{2}\\ 0\\ \sinh\frac{\phi}{2}\\ 0\end{bmatrix} \;,\quad |\psi^{(\vec{k})}_{+,\,-1}(\vec{r})\rangle = \frac{\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}}{\sqrt{\cosh\phi}}\begin{bmatrix}0\\ \cosh\frac{\phi}{2}\\ 0\\ -\sinh\frac{\phi}{2}\end{bmatrix} \;. \quad (20.181)$$

De même, pour les énergies négatives, on trouve :

$$|\psi^{(\vec{k})}_{-,\,+1}(\vec{r})\rangle = \frac{\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}}{\sqrt{\cosh\phi}}\begin{bmatrix}-\sinh\frac{\phi}{2}\\ 0\\ \cosh\frac{\phi}{2}\\ 0\end{bmatrix} \;,\quad |\psi^{(\vec{k})}_{-,\,-1}(\vec{r})\rangle = \frac{\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}}{\sqrt{\cosh\phi}}\begin{bmatrix}0\\ \sinh\frac{\phi}{2}\\ 0\\ \cosh\frac{\phi}{2}\end{bmatrix} \;. \quad (20.182)$$

Comme il se doit, les deux états $|\psi^{(\vec{k})}_{\varepsilon,\pm 1}(\vec{r})\rangle$ et $|\psi^{(\vec{k})}_{-\varepsilon,\,\pm 1}(\vec{r})\rangle$ sont orthogonaux. Pour chaque énergie, l'observable Σ_z peut prendre les deux valeurs ± 1 ; la valeur propre $\Sigma = +1$ représente un état où la composante S_z du spin est dans le même sens que l'impulsion $\hbar\vec{k}$, $\Sigma = -1$ étant associé à l'état où le spin est “en arrière” vis-à-vis de l'impulsion. Au total, on imagine une rotation autour de la direction de propagation qui est tantôt dans un sens, tantôt dans l'autre : c'est pourquoi on parle d'états à *hélicité* positive et négative, respectivement. Noter que l'hélicité peut tout autant se définir avec le moment cinétique total $\vec{J} = \vec{L} + \vec{S}$ puisque $\vec{L}.\vec{p} = 0$; on peut montrer que l'hélicité est un invariant de Lorentz.

L'identification de $\vec{\Sigma}$ avec le spin (au facteur $\frac{\hbar}{2}$ près) effectuée plus haut est d'autant plus convaincante que dans la limite non-relativiste $pc \ll mc^2$ ($\phi \ll 1$) et pour les solutions à énergie *positive* ($\varepsilon = +$), il vient (voir (20.181)) :

$$|\psi^{(\vec{k})}_{+,\,+1}(\vec{r})\rangle = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\begin{bmatrix}1\\ 0\\ \frac{\phi}{2}\\ 0\end{bmatrix} + \mathcal{O}(\phi^2) \;,\quad |\psi^{(\vec{k})}_{+,\,-1}(\vec{r})\rangle = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\begin{bmatrix}0\\ 1\\ 0\\ -\frac{\phi}{2}\end{bmatrix} + \mathcal{O}(\phi^2) \;. \quad (20.183)$$

Dans cette limite, chaque état propre de ($H_{\rm D}$, Σ_z) a une composante d'ordre zéro en $\frac{p}{mc}$ et une autre du premier ordre. Au total, à la limite stricte $c = +\infty$, la dimension de

l'espace vectoriel passe de quatre à deux[82] :

$$|\psi^{(\vec{k})}_{+,\,+1,\,c=\infty}(\vec{r})\rangle = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad , \quad |\psi^{(\vec{k})}_{+,\,-1,\,c=\infty}(\vec{r})\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad . \qquad (20.184)$$

On désigne par *petites composantes* les composantes qui tendent vers zéro dans cette limite, les autres étant naturellement appelées *grandes composantes*. Dans la limite non-relativiste, on obtient une particule d'énergie positive décrite par une fonction d'onde à *deux*-composantes, lesquelles se distinguent par la valeur propre de Σ_z, c'est-à-dire de S_z.

Le même phénomène survient pour les solutions à énergie négative (voir (20.182)) :

$$|\psi^{(\vec{k})}_{-,\,+1}(\vec{r})\rangle = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \begin{bmatrix} -\frac{\phi}{2} \\ 0 \\ 1 \\ 0 \end{bmatrix} + \mathcal{O}(\phi^2) \quad , \quad |\psi^{(\vec{k})}_{-,\,-1}(\vec{r})\rangle = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \begin{bmatrix} 0 \\ \frac{\phi}{2} \\ 0 \\ 1 \end{bmatrix} + \mathcal{O}(\phi^2) \quad ; \qquad (20.185)$$

maintenant, ce sont les troisième et quatrième composantes qui survivent à la limite $c \to +\infty$:

$$|\psi^{(\vec{k})}_{-,\,+1,\,c=\infty}(\vec{r})\rangle = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad , \quad |\psi^{(\vec{k})}_{-,\,-1,\,c=\infty}(\vec{r})\rangle = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad . \qquad (20.186)$$

Il n'existe pas d'interprétation de ces grandes composantes d'énergie négative, survivant à la limite non relativiste : c'est l'une des difficultés de la théorie de Dirac quand seules des *fonctions d'onde* sont en jeu.

En toute généralité, l'existence de solutions à énergie négative est pour le moins problématique puisque l'on s'attend à trouver que l'énergie de l'électron est bornée inférieurement par mc^2. L'interprétation proposée par Dirac est schématiquement la suivante. Tout en entérinant implicitement l'idée suivant laquelle, dans toute théorie relativiste, le nombre de particules ne saurait être fixé à cause de la possibilité de création de paires, il saute aux yeux que la proposition de Dirac repose une contradiction de fond : alors que le cadre de travail est indiscutablement une théorie à *une* particule, son interprétation exige l'introduction d'une *infinité* de particules[83].

[82] On se souvient que ce n'est pas le cas pour l'équation de Klein - Gordon, ce qui rend celle-ci inapte à décrire l'électron.

[83] La contradiction est indiscutable, puisqu'il s'agit d'une théorie qui se veut *exacte*. À l'inverse, il peut arriver que des difficultés apparemment fondamentales surgissent dans un traitement *approché*. Un premier exemple est celui de certaines théories de champ moyen, où l'aspect N-corps se cache derrière un Hamiltonien effectif décrivant des *pseudo*-particules, et où effectivement il s'avère nécessaire de redéfinir l'état fondamental. Une autre situation est celle définie par la description de l'interaction semi-classique champ-matière ; il arrive que le choix d'une jauge *particulière* s'impose pour des raisons très prosaïques – par exemple le souci de manipuler des séries rapidement convergentes (voir par exemple l'exemple célèbre de la transition de Lamb $^2\mathrm{S}_{1/2} \to {}^2\mathrm{P}_{1/2}$, discuté par Cohen-Tannoudji [81]). En pareil cas, l'objection suivant laquelle l'approximation considérée n'est pas invariante de jauge n'est pas recevable.

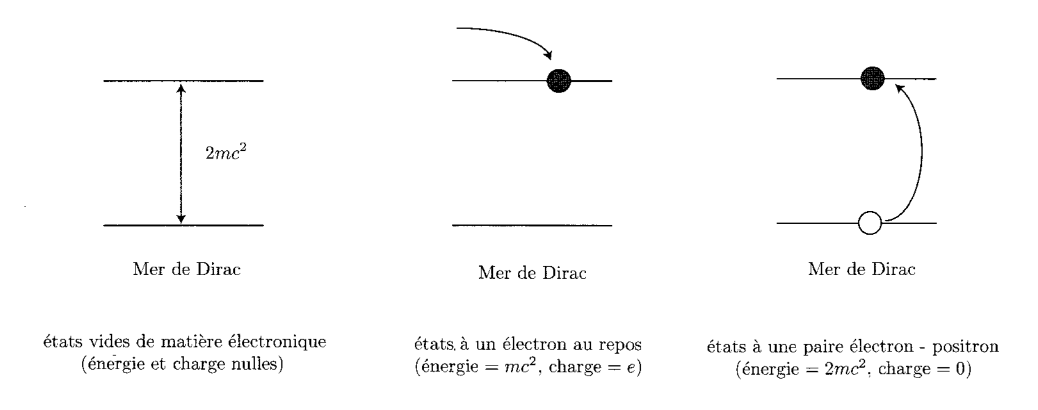


Figure 20.2: Illustration schématique de l'interprétation de Dirac.

Dirac postule en effet l'existence d'un nombre de particules occupant tous les états d'énergie négative – nombre forcément infini – chaque état à une particule situé dans cette bande pouvant être occupé par un nombre fini de particules, en pratique une seule compte tenu du spin (Principe de Pauli). Cette *bande* d'énergie remplie et saturée (on ne peut lui ajouter de particules, d'où l'importance de l'aspect *fermionique* des électrons !), est appelée *mer de Dirac* ; elle culmine en énergie à $-mc^2$ et est séparée par un *gap* égal à $2mc^2$ de l'état de plus basse énergie positive. Le système composé de la mer de Dirac remplie et d'aucune particule dans la bande positive est *défini* comme constituant le vide de matière électronique ; son énergie est conventionnellement prise égale à zéro[84], tout comme sa charge (ce qui est nettement moins intuitif), et est supposé ne donner lieu, en tant que tel, à aucune manifestation observable.

Selon Dirac, un état à un électron est formé en plaçant une particule supplémentaire dans un état d'énergie positive ; ceci coûte une énergie au moins égale à[85] $mc^2 \leq \gamma mc^2$ (il faut créer un électron) ; en disposant N particules dans N états d'énergie positive, on fabrique un système à N électrons au sens usuel. D'une façon générale, les particules situées dans la bande d'énergie positive sont interprétées comme étant les électrons "ordinaires", la mer étant pleine par ailleurs (voir fig. 20.2).

L'analyse des états d'énergie négative a de plus conduit Dirac à formuler par la suite (1931) la *Théorie des trous* (*Hole Theory*), proposant le scénario suivant. L'état vide (mer pleine) a – conventionnellement – une énergie et une charge nulles ; pour l'exciter – sans ajouter de particule de l'extérieur –, il faut lui fournir une énergie supérieure ou égale à celle du *gap*, prenant une particule d'énergie inférieure à $-mc^2$ et en la hissant dans la bande d'énergie positive[86] : ceci coûte au moins $2mc^2$. Après cette opération, on

[84]Seules les *différences d'énergie* ont un sens physique.

[85]$\gamma = \left[1 - (v/c)^2\right]^{-1/2}$.

[86]Cette image d'une infinité d'états remplis comme constituant l'état fondamental est très féconde et apparaît dans des contextes très variés, notamment dans le Problème à N-corps. Par exemple, en matière condensée de basse énergie, l'état fondamental d'un isolant ou d'un semi-conducteur intrinsèque de bande interdite E_{g} correspond à un remplissage total (par des électrons) de la bande de valence, l'équivalent de la mer négative de Dirac ; cet état fondamental, dénué de toute excitation par définition,

se retrouve avec un électron dans la bande positive et une lacune dans la mer de Dirac ; en vertu de la conservation de la charge, il faut attribuer au *trou* dans la mer de Dirac une charge positive, exactement opposée à celle de l'électron. Au départ, Dirac croyait d'ailleurs que ces lacunes positives étaient des protons[87], mais Weyl fit remarquer que ces *trous* devaient avoir exactement la même masse que l'électron (et aussi un spin 1/2). Mis tous ensemble, ces arguments permirent de conclure en prédisant l'existence du *positron*, antiparticule de l'électron (même masse, même spin, charge opposée), qui fut découvert peu après (1933) par C.D. Anderson[88]. La création d'une paire électron – positron coûte au moins deux fois l'énergie de repos $2mc^2$, soit environ 1, 2 MeV.

Ainsi, en dépit des difficultés de fond qui subsistent dans sa formulation première esquissée ci-dessus pour la particule libre – d'autres vont d'ailleurs apparaître à propos de l'atome d'hydrogène (voir juste ci-après) –, la théorie de Dirac se révèle porteuse d'une grande fécondité conceptuelle. Elle conserve d'ailleurs toute sa valeur en Théorie quantique des champs[89], le Hamiltonien de Dirac devenant un opérateur agissant sur des *champs*, et non plus sur la fonction d'onde à quatre composantes d'une particule ; par nature, les champs ainsi définis contiennent virtuellement un nombre quelconque de particules. Aussi séduisante qu'elle soit, l'interprétation de Dirac, formulée dans le cadre strict d'une théorie à une seule particule, reste fondamentalement contradictoire, mais c'est justement cette contradiction qui ouvrit la voie à la Théorie quantique des champs.

---

peut ainsi être appelé *état vide*. Les états excités se forment en prenant un électron de cette bande pour le hisser, en lui fournissant au moins l'énergie E_g (équivalent de $2mc^2$), dans la bande de conduction. De la sorte, on crée une lacune positive dans la bande de valence et une particule chargée négativement dans la bande de conduction : c'est une paire électron-trou.

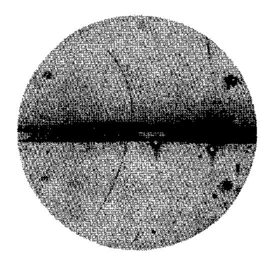
[87]On situe historiquement en 1919 la découverte du proton ($\pi\rho\tilde{\omega}\tau o\nu$ = *premier*), attribuée à Rutherford en raison de ses travaux magistraux en la matière. La masse du proton est $938,272\,013(23)$ MeV/c^2 [82].

[88]Carl David Anderson (1905-1991) a partagé le Prix Nobel en 1936 avec Victor Francis Hess (1883-1964), qui avait découvert les *rayons cosmiques*, suivant l'appellation (*cosmic rays*) introduite par Millikan en 1925. Anderson découvrit le positron en effectuant, sous la direction de ce dernier, des recherches sur le rayonnement cosmique, mettant en évidence une particule de masse comparable à celle de l'électron mais ayant une charge opposée. Les premières observations datent de 1932 ; Anderson se livra alors à un véritable travail expérimental de bénédictin pour en constituer la confirmation définitive peu après, s'autorisant alors, et alors seulement, à soumettre pour publication l'article historique de 1933 [83] – voir l'encadré p. 891.

[89]Le développement de la Théorie quantique des champs a été rendu possible par la percée de Fock définissant l'espace portant son nom (*espace de Fock*), brièvement décrit dans la note 77, p. 827.

Le résumé (*abstract*) de l'article d'Anderson[a] est le suivant :

> *Out of a group of 1300 photographs of cosmic-ray tracks in a vertical Wilson chamber 15 tracks were of positive particles which could not have a mass as great as that of the proton. From an examination of the energy-loss and ionization produced it is concluded that the charge is less than twice, and is probably exactly equal to, that of the proton. If these particles carry unit positive charge the curvatures and ionizations produced require the mass to be less than twenty times the electron mass. These particles will be called positrons. Because they occur in groups associated with other tracks it is concluded that they must be secondary particles ejected from atomic nuclei.*
>
> Editor



La photo[b] historique ci-dessus montre la trace d'un *positron* traversant une plaque de plomb de 6 mm d'épaisseur en présence d'un champ magnétique. On remarque la variation des rayons de courbure, l'énergie du positron passant de 63 à 23 MeV. Le sens de la déviation donne le *signe* de la charge, la valeur des rayons de courbure permet de déterminer la *masse* de la particule.

[a]Phys. Rev., **43**, 491 (1933).
[b]*ibid.*

20.6 Champ central et atome d'hydrogène en théorie de Dirac

Le traitement de l'atome d'hydrogène en théorie de Dirac est très riche d'enseignements ; il consiste à trouver les états propres du second membre de (20.137), avec $V(r) = -\frac{e'^2}{r}$, soit les vecteurs $|\Psi(\vec{r})\rangle$ à quatre composantes, solutions de :

$$[c\vec{\alpha}.\vec{p} + V(\vec{r})\mathbf{1}_4 + \beta mc^2]|\Psi(\vec{r})\rangle = E|\Psi(\vec{r})\rangle \ , \tag{20.187}$$

avec toujours $\vec{p} = -i\hbar\vec{\nabla}$. Mettant de côté les problèmes fonciers concernant les solutions d'énergie négative, les résultats obtenus sont remarquables et, notamment, donnent une explication presque parfaite de la structure fine de l'hydrogène, tout comme le calcul de Sommerfeld (1916) exposé dans le Tome I (sous-section 7.3.3 et problème 7.4.13) – et pour cause : la formule pour les niveaux d'énergie qui en résulte... est exactement la même que celle trouvée par Sommerfeld.

Tout comme dans le cas non-relativiste, il est très utile d'exploiter d'emblée la symétrie sphérique du potentiel afin de pouvoir séparer les variables et de n'avoir à

résoudre qu'une équation impliquant la seule variable radiale $r \stackrel{\text{déf}}{=} \|\vec{r}\|$. Ce qui suit effectue la séparation des variables radiale et angulaires et vaut pour tout potentiel central $V(r)$.

Une fois l'électron nanti des moments cinétique orbital $\vec{L}$ et de spin $\vec{S}$, une base de départ est formée par les produits tensoriels $Y_{lm} \otimes |\chi\rangle$, où χ peut être représenté par un vecteur colonne à deux éléments sur une base donnée de l'espace des états d'un spin $S = 1/2$. Choisissant les états propres de $\vec{S}^2$ et S_z notés simplement[90] $|\pm\frac{1}{2}\rangle$, $|\chi\rangle$ est de la forme :

$$|\chi\rangle = \begin{bmatrix} c_+ \\ c_- \end{bmatrix} \equiv c_+ \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_- \begin{bmatrix} 0 \\ 1 \end{bmatrix} \equiv c_+|+\frac{1}{2}\rangle + c_-|-\frac{1}{2}\rangle \ , \tag{20.188}$$

avec conventionnellement $|c_+|^2 + |c_-|^2 = 1$. Une base d'états est donc l'ensemble des vecteurs $Y_{lm_l} \otimes |m_s\rangle$ où les nombres l et m_l sont définis comme d'habitude (l entier positif ou nul, $-l \leq m_l \leq +l$, $m_s = \pm 1/2$). Cette base n'est visiblement pas adaptée puisque ni $\vec{L}$, ni $\vec{S}$ ne sont des constantes du mouvement (voir éqs. (20.161) et (20.164)). En revanche $\vec{J}$ en est une (voir (20.165)), de sorte que le bon ECOC est ici (H_{D}, $\vec{J}^2$, J_z) ; la commodité impose donc d'introduire d'emblée les vecteurs propres de la somme $\vec{J} = \vec{L}+\vec{S}$, génériquement notés $|j\, m\rangle$; les nombres j et m sont déterminés d'après les résultats connus concernant l'addition des moments cinétiques (voir chapitre 18).

Les nombres l et $s = 1/2$ étant fixés, on sait que j peut prendre les deux valeurs $j = l \pm \frac{1}{2}$ (si $l > 0$) et la seule valeur $j = 1/2$ (si $l = 0$). Les états propres $|j\, m\rangle$ s'obtiennent en combinant linéairement les $Y_{l\, m_l} \otimes |m_s\rangle$ de sorte que la somme des nombres magnétiques soit égale à m : $m_l + m_s = m$; en conséquence, seuls apparaissent les produits $Y_{l\, m\pm\frac{1}{2}} \otimes |\mp\frac{1}{2}\rangle$ dans l'expression du vecteur $|j\, m\rangle$. Utilisant la notation traditionnelle $\mathcal{Y}_l^{j,m}$ pour ces derniers, on a ainsi :

$$\mathcal{Y}_l^{l\pm\frac{1}{2},m} = C_-^{(\pm)}(l,\, m) Y_{l\, m-1/2} \otimes |+\frac{1}{2}\rangle + C_+^{(\pm)}(l,\, m) Y_{l\, m+1/2} \otimes |-\frac{1}{2}\rangle \ , \tag{20.189}$$

dont une écriture équivalente est la matrice-colonne :

$$\mathcal{Y}_l^{j=l\pm\frac{1}{2},m} = \begin{bmatrix} C_-^{(\pm)}(l,\, m) Y_{l\, m-1/2} \\ C_+^{(\pm)}(l,\, m) Y_{l\, m+1/2} \end{bmatrix} \tag{20.190}$$

quand on utilise la base propre de S_z. Les objets $\mathcal{Y}_l^{l\pm\frac{1}{2},m}$ ainsi définis sont appelés *harmoniques sphériques spinorielles* ; ce sont des spineurs à deux composantes, constituant les vecteurs propres du moment cinétique total quand le spin vaut 1/2. Il reste à trouver les coefficients $C_{\varepsilon'}^{(\varepsilon)}(l,\, m)$: c'est le sujet du problème 20.8.3. Ceci fait, on obtient :

$$\mathcal{Y}_l^{l+\frac{1}{2},m} = \sqrt{\frac{l+\frac{1}{2}+m}{2l+1}}\, Y_{l\, m-1/2} \otimes |+\frac{1}{2}\rangle + \sqrt{\frac{l+\frac{1}{2}-m}{2l+1}}\, Y_{l\, m+1/2} \otimes |-\frac{1}{2}\rangle \ , \tag{20.191}$$

[90] Il est visiblement inutile de rappeler la valeur de S dans la notation des *kets*.

$$\mathcal{Y}_l^{l-\frac{1}{2},m} = \sqrt{\frac{l+\frac{1}{2}-m}{2l+1}}\, Y_{l\,m-1/2} \otimes |+\frac{1}{2}\rangle - \sqrt{\frac{l+\frac{1}{2}+m}{2l+1}}\, Y_{l\,m+1/2} \otimes |-\frac{1}{2}\rangle \ ; \quad (20.192)$$

à nouveau, toujours sur la base propre de S_z, une écriture équivalente est :

$$\mathcal{Y}_l^{l+\frac{1}{2},m} = \begin{bmatrix} \sqrt{\frac{l+\frac{1}{2}+m}{2l+1}}\, Y_{l\,m-1/2} \\ \sqrt{\frac{l+\frac{1}{2}-m}{2l+1}}\, Y_{l\,m+1/2} \end{bmatrix} , \qquad \mathcal{Y}_l^{l-\frac{1}{2},m} = \begin{bmatrix} \sqrt{\frac{l+\frac{1}{2}-m}{2l+1}}\, Y_{l\,m-1/2} \\ -\sqrt{\frac{l+\frac{1}{2}+m}{2l+1}}\, Y_{l\,m+1/2} \end{bmatrix} . \quad (20.193)$$

Noter que l'on peut identifier $\mathcal{Y}_l^{l\pm\frac{1}{2},m} \equiv \mathcal{Y}_{j\mp\frac{1}{2}}^{j,m}$. En résumé, les $\mathcal{Y}_l^{l\pm\frac{1}{2},m}$ satisfont :

$$\vec{J}^2\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m} = \hbar^2 j(j+1)\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m} , \qquad J_z\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m} = m\hbar\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m} , \quad (20.194)$$

$$\vec{L}^2\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m} = \hbar^2(j\mp\frac{1}{2})(j\mp\frac{1}{2}+1)\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m} , \qquad \vec{S}^2\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m} = \frac{3}{4}\hbar^2\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m} . \quad (20.195)$$

Toute combinaison linéaire des deux harmoniques spinorielles $\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m}$ est donc un vecteur propre $|j\,m\rangle$ de $(\vec{J}^2,\, J_z)$. D'un autre côté, ces deux fonctions sont de parité opposée, puisque leurs nombres l diffèrent d'une unité (voir (18.80)) ; comme il s'agit d'un champ central, les solutions propres sont soit paires, soit impaires. On peut donc factoriser le problème 4×4 de Dirac en construisant une base de parité donnée, sans sortir de l'espace de j donné. Pour cela, il faut faire agir sur l'*une* des deux $\mathcal{Y}_{j\mp\frac{1}{2}}^{j,m}$ un opérateur scalaire vis-à-vis du spin et impair vis-à-vis de la parité ; le seul opérateur ayant ces propriétés est $\vec{u}.\vec{\sigma}$ où $\vec{u}$ est le vecteur unitaire porté par un vecteur $\vec{r}$ quelconque, soit $r^{-1}\vec{r}.\vec{\sigma}$. On en vient ainsi à considérer (par exemple) les deux spineurs ϕ et χ :

$$\boxed{\phi \overset{\text{déf}}{=} \mathcal{Y}_{j-\frac{1}{2}}^{j,m} \qquad \chi \overset{\text{déf}}{=} \frac{\vec{r}.\vec{\sigma}}{r}\mathcal{Y}_{j+\frac{1}{2}}^{j,m}} \quad (20.196)$$

Ils sont à la fois propres de $(\vec{J}^2,\, J_z)$, du type $|jm\rangle$, et ont une parité définie égale à $(-1)^l$ avec $l = j - \frac{1}{2}$; une combinaison linéaire *quelconque* possède les mêmes propriétés.

À ce stade, on note que, formellement, le problème 4×4 est devenu un problème 2×2 puisque l'on peut dorénavant manipuler *en bloc* chacun des deux spineurs ϕ et χ, qui sont des entités élémentaires dotées de toutes les propriétés intrinsèques souhaitables (parité, moment angulaire, un peu comme s'il s'agissait de *particules élémentaires*). On obtiendra bien *in fine* toutes les dimensions du problème 4×4 puisque, comme attendu, le spectre du Hamiltonien est impair en E : si les solutions sont écrites $F(r)\phi + G(r)\chi$ – avec les fonctions radiales solutions d'une équation à écrire... et à résoudre –, on pourra former les deux solutions propres orthogonales à quatre composantes :

$$|\Psi_+\rangle = \begin{bmatrix} F(r)\phi \\ G(r)\chi \end{bmatrix} , \qquad |\Psi_-\rangle = \begin{bmatrix} -G(r)\phi \\ F(r)\chi \end{bmatrix} ; \quad (20.197)$$

l'une correspondant à l'énergie E, l'autre à l'énergie $-E$. Dans la limite faiblement relativiste, F et G seront visiblement associées aux grandes et petites composantes, respectivement.

Introduisons maintenant l'opérateur radial p_r, comme en (19.13) :

$$p_r = \frac{1}{2}\left(\frac{\vec{r}}{r}.\vec{p} + \vec{p}.\frac{\vec{r}}{r}\right) = \frac{1}{r}(\vec{r}.\vec{p} - \mathrm{i}\hbar) \ , \qquad (20.198)$$

donné en représentation-q par (voir (19.19)) :

$$p_r = -\mathrm{i}\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right) \equiv -\mathrm{i}\hbar\frac{1}{r}\frac{\partial}{\partial r}r \ . \qquad (20.199)$$

Il est maintenant très naturel d'introduire l'opérateur radial α_r :

$$\alpha_r \stackrel{\text{déf}}{=} \frac{1}{r}\vec{r}.\vec{\alpha} \ ; \qquad (20.200)$$

de cette définition découle $\alpha_r^2 = \mathbf{1}$ puisque $\alpha_u^2 = \mathbf{1}$ quelle que soit la composante cartésienne $u = x, y, z$; de plus $\{\alpha_r, \beta\} = 0$, visiblement. Par ailleurs, compte tenu de l'expression des matrices $\vec{\alpha}$, il est facile d'établir la relation (utiliser (18.143)) :

$$(\vec{V}.\vec{\alpha})(\vec{W}.\vec{\alpha}) = \vec{V}.\vec{W}\,\mathbf{1} + \mathrm{i}\vec{\Sigma}.(\vec{V}\times\vec{W}) \ , \qquad (20.201)$$

où $\vec{V}$ et $\vec{W}$ sont deux vecteurs commutant avec $\vec{\alpha}$; on a ainsi :

$$(\vec{r}.\vec{\alpha})(\vec{p}.\vec{\alpha}) = \vec{r}.\vec{p}\,\mathbf{1} + \mathrm{i}\vec{\Sigma}.(\vec{r}\times\vec{p}) \ ; \qquad (20.202)$$

le premier membre est $r\alpha_r\,\vec{p}.\vec{\alpha}$; multipliant membre à membre à gauche par α_r, il vient :

$$r\,\vec{p}.\vec{\alpha} = \vec{r}.\vec{p}\,\alpha_r + \mathrm{i}\alpha_r\,\vec{\Sigma}.(\vec{r}\times\vec{p}) \quad \Longleftrightarrow \quad \vec{p}.\vec{\alpha} = \frac{1}{r}\,\vec{r}.\vec{p}\,\alpha_r + \frac{\mathrm{i}}{r}\alpha_r\,\vec{\Sigma}.\vec{L} \ ; \qquad (20.203)$$

comme $r^{-1}\vec{r}.\vec{p} = p_r + \mathrm{i}\hbar r^{-1}$ (voir (20.198)), l'expression de la partie cinétique de H_{D} est :

$$c\vec{p}.\vec{\alpha} = c\alpha_r\left[p_r + \frac{\mathrm{i}}{r}\left(\hbar + \frac{2}{\hbar}\vec{L}.\vec{S}\right)\right] \ , \qquad (20.204)$$

qui met en lumière l'interaction spin-orbite, proportionnelle à $2\vec{L}.\vec{S} = \vec{J}^2 - \vec{L}^2 - \vec{S}^2$. Compte tenu de ces résultats, le Hamiltonien de Dirac pour un champ à symétrie sphérique prend la forme :

$$\boxed{H_{\mathrm{D}} = c\alpha_r\left(-\mathrm{i}\hbar\frac{\partial}{\partial r} + \frac{2\mathrm{i}}{\hbar r}\,\vec{L}.\vec{S}\right) + V(r)\,\mathbf{1} + mc^2\beta} \qquad (20.205)$$

Une autre forme très utile [53] est rendue possible par l'introduction de l'opérateur noté Q :

$$\mathsf{Q} \stackrel{\text{déf}}{=} \beta(\hbar^{-1}\vec{\Sigma}.\vec{L} + \mathbf{1}) \equiv \beta(2\hbar^{-2}\vec{S}.\vec{L} + \mathbf{1}) \ ; \qquad (20.206)$$

il est facile de voir que le Hamiltonien H_{D} (20.205) est[91] :

$$\boxed{H_{\mathrm{D}} = c\alpha_r\left(p_r + \frac{\mathrm{i}\hbar}{r}\beta\mathsf{Q}\right) + V(r)\,\mathbf{1} + mc^2\beta} \qquad (20.207)$$

[91] Prendre garde à la définition de H_{D} utilisée par Schiff [53], où les signes de α et β sont inversés (voir son éq. (44.1).

qui est une expression alternative pour un champ central. On voit sans difficulté que $[\alpha_r, \mathsf{Q}] = [\beta, \mathsf{Q}] = [p_r, \mathsf{Q}] = 0$, de sorte de $[\mathsf{Q}, H] = 0$. Les valeurs propres de K sont remarquablement simples ; en effet, on a, comme $[\beta, \vec{\Sigma}] = 0$ et que $\beta^2 = \mathbf{1}$, on a :

$$\hbar^2 \mathsf{Q}^2 = (\vec{\Sigma}.\vec{L})^2 + 2\hbar\vec{\Sigma}.\vec{L} + \hbar^2\mathbf{1} \ . \tag{20.208}$$

En utilisant (18.143), on a $(\vec{\Sigma}.\vec{L})^2 = \vec{L}^{\,2} + \mathrm{i}\vec{\Sigma}.(\vec{L}\times\vec{L})$; avec $\vec{L}\times\vec{L} = \mathrm{i}\hbar\vec{L}$, il vient :

$$\hbar^2 \mathsf{Q}^2 = \vec{L}^{\,2} + \hbar\vec{\Sigma}.\vec{L} + \hbar^2\mathbf{1} = (\vec{L} + \frac{\hbar}{2}\vec{\Sigma})^2 + \frac{\hbar^2}{4}\mathbf{1} \ , \tag{20.209}$$

puisque $\vec{\Sigma}^2 = 3\,\mathbf{1}$. On reconnaît le moment cinétique total $\vec{J} = \vec{L} + \vec{S}$, d'où :

$$\hbar^2 \mathsf{Q}^2 = \vec{J}^2 + \frac{\hbar^2}{4}\mathbf{1} \ . \tag{20.210}$$

Les valeurs propres de Q^2 sont donc $j(j+1) + \frac{1}{4} \equiv (j + \frac{1}{2})^2$. Comme j est un demi-entier positif, les valeurs propres de Q sont les entiers relatifs sauf zéro. Dans la suite, on prend une représentation où H_{D} et Q sont simultanément diagonaux, notant $\mathsf{q} \equiv \pm(j + \frac{1}{2}) \in \mathbb{Z}^*$ la valeur propre de Q, E celle de H_{D}.

Afin d'écrire explicitement les équations radiales, il convient de choisir une représentation pour les deux opérateurs α_r et β, dans l'espace de dimension 2 des spineurs ϕ et χ. Il n'apparaît plus maintenant que *deux* opérateurs de Dirac, α_r et β, qui doivent satisfaire :

$$\alpha_r^2 = \beta^2 = \mathbf{1} \ , \qquad \alpha_r\beta + \beta\alpha_r = 0 \ . \tag{20.211}$$

Fort heureusement, ces deux seules relations peuvent être respectées par des matrices de dimension 2 : la diminution du nombre d'opérateurs de Dirac, et celle corrélative du nombre de relations de commutation à satisfaire, est compatible avec la réduction de la dimension de l'espace de travail obtenue grâce à l'exploitation de la symétrie. La représentation matricielle des deux opérateurs doit impérativement respecter les relations (20.211) (la représentation doit être *fidèle*) ; cela étant assuré, le choix est libre et reste une pure affaire de convenance. Visiblement, on peut prendre pour α_r et β deux quelconques des trois matrices de Pauli ; la présence du facteur[92] i dans le terme en α_r incite à prendre σ_y pour représenter α_r ; dès lors, il est commode de prendre σ_z pour représenter β. Ce choix étant fait, le Hamiltonien (20.207) est représenté dans l'espace des spineurs par la *matrice* 2×2 notée $M_{H_{\mathrm{D}}}$:

$$M_{H_{\mathrm{D}}} = c\sigma_y\big(p_r + \frac{\mathrm{i}\hbar}{r}\,\sigma_z\mathsf{q}\big) + V(r)\,\mathbf{1}_2 + mc^2\sigma_z \qquad (\mathsf{q}\in\mathbb{Z}^*) \ , \tag{20.212}$$

les σ_u étant les matrices de Pauli conventionnelles (dans leur représentation standard). En désignant par F et G les deux fonctions radiales à trouver, l'équation à résoudre s'écrit explicitement[93] :

$$\boxed{\begin{bmatrix} V(r) + mc^2 & \hbar c\left(-\frac{1}{r}\frac{\partial}{\partial r}r - \frac{\mathsf{q}}{r}\right) \\ \hbar c\left(\frac{1}{r}\frac{\partial}{\partial r}r - \frac{\mathsf{q}}{r}\right) & V(r) - mc^2 \end{bmatrix}\begin{bmatrix} F(r) \\ G(r)\end{bmatrix} = E\begin{bmatrix} F(r) \\ G(r)\end{bmatrix}} \tag{20.213}$$

[92] p_r apporte également un facteur i.

[93] On a utilisé au passage $\sigma_y\sigma_z = \mathrm{i}\sigma_x$.

Il s'agit d'un système différentiel du premier ordre, donnant donc deux équations couplées ; si, comme en théorie non-relativiste, on pose $F(r) = r^{-1}u(r)$, $G(r) = r^{-1}v(r)$, on obtient :

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}r}\begin{bmatrix} u(r) \\ v(r) \end{bmatrix} = \begin{bmatrix} \frac{\mathsf{q}}{r} & \frac{1}{\hbar c}(mc^2 + E - V) \\ \frac{1}{\hbar c}(mc^2 + V - E) & -\frac{\mathsf{q}}{r} \end{bmatrix}\begin{bmatrix} u(r) \\ v(r) \end{bmatrix}} \tag{20.214}$$

Ici s'achève la séparation des variables pour un champ central $V(r)$ quelconque. Un choix ou un autre de la fonction $V(r)$ fixe cette équation[94] ; la résolution de celle-ci, complétée avec les bonnes conditions aux limites, fournit les fonctions radiales cherchées. En particulier, en cas de saut fini de potentiel, on est conduit à écrire à nouveau des équations de raccordement, exprimant la continuité des *seules* fonctions, nullement de leurs dérivées premières respectives ; si on veut l'imposer, on trouve seulement la solution... identiquement nulle : l'équation différentielle est du *premier* ordre en espace, alors qu'elle du second ordre (par le Laplacien) en théorie non-relativiste. C'est donc ici la dérivée *première* qui encaisse les sauts de potentiel, quand ils existent. Une fois les fonctions obtenues $F(r)$ et $G(r)$, on peut reconstruire les solutions à quatre composantes suivant la prescription (20.197).

En ce qui concerne le potentiel Coulombien, $V(r) = -\frac{e'^2}{r}$, le calcul a fait l'objet de nombreux travaux, depuis les articles de Dirac [76], Darwin[95] [86] et Gordon [87] jusqu'à l'approche récente de Martinez-y-Romero [88], en passant par le traitement classique de Bethe et Salpeter [89], ou de Bjorken et Drell [84]. Dans la présentation de Schiff, la suite du calcul est relativement routinière (voir [53] pour les détails) : on cherche les fonctions $F(r)$ et $G(r)$ sous la forme de séries entières, multipliées par le monôme r^ν, ν étant un exposant à trouver ; comme pour la théorie non-relativiste, la condition aux limites en $r = 0$ est toujours (voir (19.40)) :

$$\lim_{r\to 0} rF = \lim_{r\to 0} rG = 0 \ , \tag{20.215}$$

ce qui impose $\nu > -1$ (ou plus généralement $\Re\nu > -1$, voir ci-dessous...). L'analyse de ces séries couplées montre, une fois encore, qu'elles doivent se réduire à des polynômes pour que les fonctions aient les bonnes propriétés à l'infini. Une particularité du calcul relativiste est le fait que l'exposant ν n'est pas entier ; on trouve :

$$\nu = \sqrt{\mathsf{q}^2 - \alpha^2} - 1 \ , \tag{20.216}$$

où α désigne ici la constante de structure fine (l'exposant redevient entier à la limite $c = \infty$). Le calcul *à la Dirac* met ainsi en évidence un comportement singulier à l'origine de la fonction propre fondamentale ($j = 1/2, \mathsf{q} = 1$), puisque l'exposant ν est *négatif*, quoique très petit en valeur absolue (il vaut environ $-\alpha^2/2$) : la divergence est "douce"

[94] À nouveau, la limite du puits en fonction de Dirac est délicate [42].

[95] Dès 1913, en rapport avec la diffusion Rutherford, Darwin avait mis en évidence [85], dans un cadre strictement classique, des trajectoires en spirale spécifiques du traitement *relativiste* du problème de Kepler.

(elle est absente[96] si, strictement, $\alpha = 0$ – limite non-relativiste prise à $\hbar$ fixée). Noter toutefois que les fonctions $rF(r)$ et $rG(r)$ se comportent comme $r^{\sqrt{\mathsf{q}^2-\alpha^2}}$, et satisfont la condition aux limites (20.215) tant que la racine carrée est *réelle.*

Le calcul complet pour les états d'énergie positive, donne l'énergie propre sous la forme d'une fonction du nombre k restreint aux valeurs positives[97] ($\mathsf{q} = j + 1/2$) et d'un autre nombre $n' \in \mathbb{N}$. En posant $n = n' + (j + 1/2)$, on obtient l'écriture traditionnelle :

$$\boxed{E_{nj} = \frac{mc^2}{\sqrt{1 + \frac{\alpha^2}{\left(n-j-\frac{1}{2}+\sqrt{(j+\frac{1}{2})^2-\alpha^2}\right)^2}}} \equiv E(n,\, j) + mc^2} \tag{20.217}$$

avec $n \in \mathbb{N}^*$, j demi-entier $\leq n - \frac{1}{2}$. L'expression de $E(n,\, j)$ est à comparer avec celle obtenue par Sommerfeld (Tome I, p. 220) : les deux sont strictement identiques, pourvu que l'on accepte l'identification $n_\theta \leftrightarrow j + \frac{1}{2}$... Pour l'état fondamental $n = 1$, $l = 0$, $j = 1/2$, la formule (20.217) donne :

$$E(1,\, 1/2) = mc^2\,\sqrt{1 - \alpha^2} - mc^2 \simeq -\frac{\alpha^2}{2}\, mc^2 \ , \tag{20.218}$$

en conformité avec (19.173). Les états liés d'énergie positive sont donc situés entre $mc^2\sqrt{1-\alpha^2}$ et mc^2. Les états "liés" d'énergie négative sont situés symétriquement par rapport à $E = 0$ mais, tout comme précédemment, ne peuvent recevoir d'interprétation satisfaisante.

D'autres difficultés surgissent quand on considère un atome hydrogénoïde, pour lequel il suffit de remplacer partout e'^2 par Ze'^2, ce qui revient à effectuer la substitution $\alpha \rightarrow Z\alpha$ dans les expressions précédentes ; l'exposant ν est alors égal à $\sqrt{\mathsf{q}^2 - (Z\alpha)^2} - 1$. Pour l'état fondamental ($\mathsf{q} = 1$), ν devient complexe pour $Z \gtrsim 137$, un fait à considérer comme un nouveau symptôme grave, puisque l'on ne peut plus satisfaire les conditions aux limites à l'origine pour les fonctions radiales – d'ailleurs, l'énergie elle-même devient complexe , ce qui, en bonne logique, signifie que cet état disparaît en tant que solution acceptable[98]. Passant outre cette anomalie pour creuser la difficulté, on voit que les fonctions radiales rF et rG, qui n'ont plus de limite quand $r \rightarrow 0$, oscillent violemment sur une échelle de longueur[99] $\lesssim \alpha a_0 \lesssim 400$ F ; il n'existe aucune interprétation physique

[96] Au passage, on note que les deux limites $\hbar \rightarrow 0$ et $c \rightarrow +\infty$ ne commutent pas puisque les deux constantes $\hbar$ et c se multiplient l'une l'autre dans la combinaison globale α, qui tend ainsi vers zéro si la limite $c \rightarrow \infty$ est prise en premier, vers l'infini si on fait d'abord $\hbar \rightarrow 0$.

[97] En fait, seule la quantité k^2 intervient dans le calcul ; on a bien l'ensemble des solutions en choisissant k d'un signe donné.

[98] Ceci peut d'ailleurs être rapproché d'un phénomène lui aussi spécifique du calcul *relativiste* de Sommerfeld (Tome I, note 47 p. 219), montrant que si le moment cinétique est plus petit que $Z\alpha\hbar$, il n'existe plus de mouvement lié puisqu'il y a *capture* de l'électron par le noyau.

Par ailleurs, on note la simultanéité de deux faits : l'impossibilité de satisfaire les conditions aux limites et l'apparition d'une valeur propre complexe pour le Hamiltonien. Tout ceci est cohérent : on sait bien que l'hermiticité d'un opérateur est conditionnée par la satisfaction de certaines conditions aux limites, l'exemple le plus simple étant fourni par l'opérateur $p_x = -\mathrm{i}\hbar\partial_x$.

[99] Avec $A = 2Z$, la formule phénoménologique pour le rayon de charge R_c du noyau (Tome I, éq. (8.17)) donne $R_\mathrm{c} \sim 10\,\mathrm{F}$: l'anomalie de comportement se produit donc bien avant de s'approcher de l'espace nucléaire.

de ce fait [84] : il y a bien persistence de problèmes graves dès que les solutions d'énergie négative sont réellement impliquées.

La limite faiblement relativiste fait encore apparaître les deux types de composantes [84], grandes et petites, déjà rencontrées à propos de la particule libre ; à la limite $c = \infty$, les deux grandes composantes d'énergie positive coïncident avec les fonctions propres de l'atome d'hydrogène $\psi_{nlm}(\vec{r})$ nanties comme il se doit du spineur associé à $S = 1/2$, les deux petites composantes étant nulles, bien entendu – mais les deux grandes composantes d'énergie *négative* restent tout autant énigmatiques...

Malgré ces difficultés de fond, la théorie de Dirac – tronquée aux seules solutions d'énergie positive – reproduit avec une extraordinaire précision le spectre de l'atome d'hydrogène, un succès reposant sur le fait que seules des énergies très petites devant mc^2 sont pertinentes en la matière. En adoptant la notation spectroscopique $n\mathcal{L}_j$, où $\mathcal{L} =$ S, P, D, ... selon que $l = 0, 1, 2, ...$, l'expression (20.217) pour $E_{nj} \equiv E(n\mathcal{L}_j)$ donne :

$$E(1\mathrm{S}_{1/2}) = mc^2\sqrt{1-\alpha^2} \ , \qquad E(2\mathrm{S}_{1/2}) = E(2\mathrm{P}_{1/2}) = mc^2\sqrt{\frac{1+\sqrt{1-\alpha^2}}{2}} \ , \quad (20.219)$$

$$E(2\mathrm{P}_{3/2}) = mc^2\sqrt{1-\alpha^2/4} \ ; \quad (20.220)$$

L'écart entre les deux états $2\mathrm{P}_{1/2}$ et $2\mathrm{P}_{3/2}$ (dégénérés dans la limite non-relativiste, où j n'apparaît pas) est un exemple de *splitting* donnant lieu à la structure fine, engendrant le doublet H_α de la raie la plus énergétique de la série de Balmer. Par ailleurs, noter la dégénérescence des deux niveaux $2\mathrm{S}_{1/2}$ et $2\mathrm{P}_{1/2}$.

Jusqu'en 1947, les résultats obtenus dans la théorie de Dirac étaient en remarquable accord avec les spectres observés, date à laquelle Lamb et Retherford découvrirent, grâce à des mesures très précises, que ces deux derniers niveaux avaient des énergies différentes. L'explication ne peut être fournie qu'en dépassant la théorie de Dirac, et en quantifiant le rayonnement électromagnétique. On découvre ainsi le rôle des *fluctuations du vide* du champ : quand il est dans son état fondamental (nombre de photons nul *en moyenne*, champs $\vec{\mathcal{E}}$ et $\vec{\mathcal{B}}$ nuls *en moyenne*), le champ *quantifié* possède des fluctuations – tout comme un oscillateur peut avoir $\langle x\rangle = 0$ mais avec, forcément, $\Delta x^2 > 0$ (Heisenberg !). Les fluctuations du nombre de photons autour d'une valeur moyenne nulle se visualisent comme étant des émissions/absorption (virtuelles) de photon par l'atome, provoquant inévitablement des déplacements d'énergie[100] pour les niveaux atomiques "nus" ; un tel déplacement est généralement appelé[101] *Lamb shift.* Si pour l'état $2\mathrm{P}_{1/2}$,

[100]On doit à Bethe la première explication de ce déplacement d'énergie ; le calcul de Bethe peut être considéré comme la pose de la première pierre de l'Électrodynamique quantique. Pour une explication élémentaire, voir chapitre 25, p. 1167.

[101]Willis Eugene Lamb, Jr. (1913-) a partagé le Nobel en 1955 avec Polykarp Kusch *"for his discoveries concerning the fine structure of the hydrogen spectrum"* ; la mention pour Kusch est *"for his precision determination of the magnetic moment of the electron"*.

Lamb rata de peu [90], 19 ans avant Mössbauer, la découverte de l'effet portant le nom de ce dernier, phénomène d'apparence paradoxale (*canon sans recul*), lié au fait que les vibrations d'un solide sont *quantifiées*. Les conséquences physiques sont importantes : alors que la résonance est impossible pour un noyau libre au contraire d'un atome (voir Tome I, section 6.3.1, problème 6.4.6 et section 8.5, p. 253), il est possible de l'observer quand le noyau instable est inséré dans une matrice cristalline.

l'effet est numériquement très faible, il est au contraire notable pour l'état $2S_{1/2}$, qui est légèrement déstabilisé. Au total, ces deux états ne sont plus dégénérés, l'écart étant de l'ordre de 1 060 MHz. Oubliant l'impossibilité de la théorie de Dirac à prédire cet effet, on ne peut que constater l'extraordinaire succès de cette formalisation concernant le spectre de l'hydrogène ; il tient, bien évidemment, au fait que dans le domaine d'énergie de l'eV, il s'agit d'un cadre théorique parfaitement recevable pourvu que l'on délaisse toutes les difficultés conceptuelles liées aux états d'énergie négative.

Les quelques situations examinées dans cette section montrent bien que, telle qu'elle est formulée à ce stade (une théorie à *une* particule), l'équation de Dirac possède de graves défauts, qui deviennent absolument rédhibitoires aux hautes énergies. L'impasse dans laquelle elle a semblé conduire a permis l'émergence de la Théorie quantique des champs, où le problème des énergies négatives ne se pose plus puisqu'elle est fondamentalement une théorie où le nombre de "particules" n'est pas fixé – d'où l'importance de la notion d'espace de Fock – et où une "particule" apparaît comme l'excitation d'un certain champ (quantifié). Dans un tel cadre, il existe un état fondamental (le *vide*) et, par nature, toutes les énergies sont des excitations au-dessus de cette référence : toute énergie est donc essentiellement *positive* ; pour le champ de Dirac, il existe un *gap* d'énergie entre l'état fondamental et les premières excitations, *gap* égal à mc^2, qui est donc l'énergie de création d'*une* particule ; comme cette particule peut avoir une charge $-|e|$ ou $+|e|$, il s'agit soit d'un électron, soit d'un positron.

20.7 Limite faiblement relativiste et Hamiltonien de Pauli

Il est utile d'examiner un peu plus en détail la limite faiblement relativiste de l'équation de Dirac, en ignorant délibérément tout ce qui a trait aux solutions d'énergie négative. Le succès constaté pour l'hydrogène légitime que cette limite soit formalisée. D'une part, ceci permet de renforcer l'interprétation de l'observable $\vec{\Sigma}$ comme ingrédient du spin de l'électron, d'autre part (mêmes causes, mêmes effets), elle fournit un Hamiltonien effectif capable de décrire les phénomènes fins de la Physique atomique, où les corrections relativistes jouent un rôle et sont parfaitement observables.

Explicitons d'abord l'autre justification du fait que l'opérateur $\vec{\Sigma}$ (20.156) apporte le degré de spin, suivant la relation précise (20.166). En présence d'un champ électromagnétique $\vec{\mathcal{E}}, \vec{\mathcal{B}}$ dérivant du potentiel $(\vec{A},\, \phi)$, l'équation de Dirac est :

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(\vec{r},\, t) = [c\vec{\alpha}.(\vec{p} - e\vec{A}) + e\phi\mathbf{1}_4 + \beta mc^2]\Psi(\vec{r},\, t) \qquad (20.221)$$

ce que l'on peut écrire (omettant partout la matrice identité $\mathbf{1}_4$) :

$$(\mathcal{H} - c\vec{\alpha}.\vec{\varpi} + \beta mc^2)\Psi(\vec{r},\, t) = 0 \ , \qquad (20.222)$$

où les opérateurs $\mathcal{H}$ et $\vec{\varpi}$ sont définis comme :

$$\mathcal{H} \stackrel{\text{déf}}{=} \mathrm{i}\hbar\frac{\partial}{\partial t} - e\phi \ , \qquad \vec{\varpi} \stackrel{\text{déf}}{=} (\vec{p} - e\vec{A}) \ . \tag{20.223}$$

Multiplions maintenant (20.222) à gauche par $(\mathcal{H} + c\vec{\alpha}.\vec{\varpi} - \beta mc^2)$. Compte tenu de $\alpha_u\beta + \beta\alpha_u = 0$ (voir (20.144)) et en laissant à gauche les termes carrés, il vient :

$$\left(\mathcal{H}^2 - c^2(\vec{\alpha}.\vec{\varpi})^2 - m^2c^4\right)\Psi = c[\mathcal{H},\, \vec{\alpha}.\vec{\varpi}]\,\Psi \ . \tag{20.224}$$

Utilisant (20.201) et prenant en compte la relation $\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$, on voit que le premier membre de (20.224) est égal à :

$$\left(\mathcal{H}^2 - c^2\vec{\varpi}^2 - m^2c^4 + e\hbar c^2\vec{\Sigma}.\vec{\mathcal{B}}\right)\Psi \ . \tag{20.225}$$

Le commutateur au second membre de (20.224), $[\mathrm{i}\hbar\frac{\partial}{\partial t} - e\phi,\, \vec{\alpha}.(\vec{p} - e\vec{A})]$, donne successivement les deux termes :

$$[\mathrm{i}\hbar\frac{\partial}{\partial t},\, \vec{\alpha}.(\vec{p} - e\vec{A})] = -\mathrm{i}e\hbar\vec{\alpha}.\frac{\partial\vec{A}}{\partial t} \ , \qquad -e[\phi,\, \vec{\alpha}.(\vec{p} - e\vec{A})] = -\mathrm{i}e\hbar\vec{\alpha}.\vec{\nabla}\phi \ ; \tag{20.226}$$

compte tenu de $\vec{\mathcal{E}} = -\frac{\partial\vec{A}}{\partial t} - \vec{\nabla}\phi$, le commutateur est donc égal à $\mathrm{i}e\hbar\vec{\alpha}.\vec{\mathcal{E}}$. En définitive, (20.224) se récrit :

$$\left(\mathcal{H}^2 - m^2c^4 - c^2\vec{\varpi}^2 + e\hbar c^2\vec{\Sigma}.\vec{\mathcal{B}} - \mathrm{i}e\hbar c\vec{\alpha}.\vec{\mathcal{E}}\right)\Psi = 0 \ . \tag{20.227}$$

Pour un état stationnaire $\Psi = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Et}\,\psi$, le terme $\mathrm{i}\hbar\frac{\partial}{\partial t}$ donne simplement $E\Psi$. Après simplification du facteur temporel, (20.228) devient :

$$[(E - e\phi)^2 - m^2c^4]\psi = \left[c^2(\vec{p} - e\vec{A})^2 - e\hbar c^2\vec{\Sigma}.\vec{\mathcal{B}} + \mathrm{i}e\hbar c\vec{\alpha}.\vec{\mathcal{E}}\right]\psi \ . \tag{20.228}$$

Jusqu'à présent, aucune approximation n'a été faite. Examinons maintenant la limite faiblement relativiste en posant :

$$E = E_{\text{fr}} + mc^2 \ , \tag{20.229}$$

avec l'hypothèse supplémentaire $E_{\text{fr}},\, e\phi \ll mc^2$. Les termes du crochet au premier membre de (20.228) se simplifient en :

$$(E - e\phi)^2 - m^2c^4 \simeq 2mc^2(E_{\text{fr}} - e\phi) \ , \tag{20.230}$$

ce qui permet de récrire (20.228) comme suit :

$$E_{\text{fr}}\psi = \left[\frac{1}{2m}(\vec{p} - e\vec{A})^2 + e\phi - \frac{e\hbar}{2m}\vec{\Sigma}.\vec{\mathcal{B}} + \mathrm{i}\frac{e\hbar}{2mc}\vec{\alpha}.\vec{\mathcal{E}}\right]\psi \ , \tag{20.231}$$

soit, en rétablissant le temps "à l'envers"[102] :

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} = \left[\frac{1}{2m}(\vec{p} - e\vec{A})^2 + e\phi - \frac{e\hbar}{2m}\vec{\Sigma}.\vec{\mathcal{B}} + \mathrm{i}\frac{e\hbar}{2mc}\vec{\alpha}.\vec{\mathcal{E}}\right]\Psi \ . \tag{20.232}$$

[102]Ceci rappelle la procédure suivie par Schrödinger pour établir son équation, voir Tome I, p. 284.

Ψ est toujours un vecteur à 4 composantes. Comme il s'agit maintenant de la limite faiblement relativiste et compte tenu de l'approximation déjà effectuée, il suffit de considérer les deux premières composantes (les *grandes* composantes) – étant entendu que l'on ne considère toujours que les solutions d'énergie positive. En désignant par Ψ_+ la projection de Ψ dans le sous-espace 2×2 de ces grandes composantes, et puisque $\vec{\alpha}$ mélange les grandes et les petites composantes (celles-ci étant d'ordre v/c par rapport au premières), l'équation (20.232) devient à cet ordre :

$$\mathrm{i}\hbar\frac{\partial\Psi_+}{\partial t} = \left[\frac{1}{2m}(\vec{p} - e\vec{A})^2 + e\phi - \frac{e\hbar}{2m}\vec{\sigma}.\vec{\mathcal{B}}\right]\Psi_+ \ , \tag{20.233}$$

où $\vec{\sigma}$ est maintenant l'ensemble des trois matrices de Pauli (le terme $\mathrm{i}(e\hbar)/(2mc)\vec{\alpha}.\vec{\mathcal{E}}$ ainsi négligé produit l'interaction spin-orbite, voir ci-dessous). Les deux premiers termes du second membre sont familiers ; le troisième est un couplage du genre $-\vec{\mu}_{\mathrm{S}}.\vec{\mathcal{B}}$ et permet bien d'identifier le moment magnétique lié au spin en posant :

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma} \ , \tag{20.234}$$

justifiant à nouveau l'interprétation de $\vec{\Sigma}$, et tirant profit du fait que dans le sous-espace de projection (réduit aux deux grandes composantes), $\vec{\Sigma}$ peut être effectivement remplacé par $\vec{\sigma}$.

Compte tenu de (20.234), le Hamiltonien au second membre de (20.233) prend la forme :

$$\boxed{H_{\mathrm{Pauli}} \stackrel{\text{déf}}{=} \frac{1}{2m}(\vec{p} - e\vec{A})^2 + e\phi - \frac{e}{m}\vec{S}.\vec{\mathcal{B}}} \tag{20.235}$$

et s'appelle le Hamiltonien de Pauli. Ce résultat justifie la procédure pragmatique utilisée en théorie franchement non-relativiste, qui consiste à prendre le Hamiltonien ordinaire pour une particule chargée dans un champ électromagnétique et à rajouter *à la main* le couplage magnétique $-\vec{\mu}_{\mathrm{S}}.\vec{\mathcal{B}}$ résultant du spin, avec g_{e} strictement égal à 2.

Enfin, montrons comment obtenir l'expression correcte de l'opérateur de couplage spin-orbite V_{magn}, donné en (20.6) ou (20.9), après correction du facteur 1/2 provenant de la précession de Thomas. En repartant du Hamiltonien de Dirac, en l'absence de champ extérieur mais pour une particule liée par le potentiel à symétrie sphérique $V(r)$, on a, pour un état stationnaire ψ d'énergie E :

$$E\psi = \left[c\vec{\alpha}.\vec{p} + \beta mc^2 + V(r)\right]\psi \ . \tag{20.236}$$

En désignant par ψ_+ et ψ_- respectivement les deux doublets formés des composantes (ψ_1, ψ_2) et (ψ_3, ψ_4), l'équation (20.236) s'explicite en deux équations 2×2 couplées :

$$E\psi_+ = (mc^2 + V)\mathbf{1}_2\psi_+ + c\vec{\sigma}.\vec{p}\psi_- \ , \tag{20.237}$$

$$E\psi_- = (-mc^2 + V)\mathbf{1}_2\psi_- + c\vec{\sigma}.\vec{p}\psi_+ \ ; \tag{20.238}$$

la deuxième équation donne $(E + mc^2 - V)\mathbf{1}_2\psi_- = c\vec{\sigma}.\vec{p}\psi_+$, d'où, par report dans (20.237) :

$$(E - mc^2 - V)\mathbf{1}_2\psi_+ = \frac{1}{m}\vec{\sigma}.\vec{p}\frac{1}{1 + \frac{E-V}{mc^2}}\vec{\sigma}.\vec{p}\psi_+ \ , \tag{20.239}$$

ou encore, posant à nouveau $E = E_{\rm fr} + mc^2$:

$$(E_{\rm fr} - V)\mathbf{1}_2\psi_+ = \frac{1}{2m}\vec{\sigma}.\vec{p}\,\frac{1}{1+\frac{E_{\rm fr}-V}{2mc^2}}\vec{\sigma}.\vec{p}\,\psi_+ \ . \tag{20.240}$$

Pour expliciter le second membre, on écrit $\vec{\sigma}.\vec{p}f(\vec{r}) = [\vec{\sigma}.\vec{p},\, f(\vec{r})] + f(\vec{r})\vec{\sigma}.\vec{p}$, $f(\vec{r})$ étant une fonction quelconque. Le commutateur vaut $-\mathrm{i}\hbar\vec{\sigma}.\vec{\nabla}f(\vec{r})$, d'où :

$$\vec{\sigma}.\vec{p}f(\vec{r})\vec{\sigma}.\vec{p} = -\mathrm{i}\hbar\vec{\sigma}.(\vec{\nabla}f(\vec{r}))\,\vec{\sigma}.\vec{p} + f(\vec{r})(\vec{\sigma}.\vec{p})^2 \ . \tag{20.241}$$

En utilisant (18.143) et $\vec{p}\times\vec{p} = 0$ puisque $[p_u,\, p_v] = 0$ quels que soient u et v, il vient :

$$\vec{\sigma}.\vec{p}f(\vec{r})\vec{\sigma}.\vec{p} = -\mathrm{i}\hbar[(\vec{\nabla}f(\vec{r})).\vec{p}\,\mathbf{1}_2 + \mathrm{i}\vec{\sigma}.(\vec{\nabla}f(\vec{r}))\times\vec{p}] + f(\vec{r})\vec{p}^{\,2}\,\mathbf{1}_2 \ . \tag{20.242}$$

La fonction $f(\vec{r})$ est ici la quantité $(1+\frac{E_{\rm fr}-V}{2mc^2})^{-1}$, très proche de $1-\frac{E_{\rm fr}-V}{2mc^2}$ dans la limite faiblement relativiste, et alors $\vec{\nabla}f(\vec{r}) \simeq \frac{1}{2mc^2}\vec{\nabla}V$; prenant V à symétrie sphérique, son gradient est $\vec{\nabla}V = r^{-1}V'(r)\vec{r}$. Dans ces conditions, (20.240) devient alors (noter que $\vec{\nabla}V.\vec{p} = r^{-1}V'\vec{r}.\vec{p} = -\mathrm{i}\hbar r^{-1}V'r\partial_r = -\mathrm{i}\hbar V'\,\partial_r$) :

$$2m(E_{\rm fr} - V)\psi_+ = \left[\left(1 - \frac{E_{\rm fr}-V}{2mc^2}\right)\vec{p}^{\,2}\,\mathbf{1}_2 - \frac{\hbar^2}{2mc^2}V'\frac{\partial}{\partial r} + \frac{\hbar}{2mc^2}\frac{1}{r}V'\,(\vec{r}\times\vec{p}).\vec{\sigma}\right]\psi_+ \ . \tag{20.243}$$

On reconnaît le moment cinétique $\vec{L} = \vec{r}\times\vec{p}$, d'où, puisque $\vec{S} = \hbar\vec{\sigma}/2$:

$$E_{\rm fr}\psi_+ = \left[\frac{1}{2m}\left(1 - \frac{E_{\rm fr}-V}{2mc^2}\right)\vec{p}^{\,2} - \frac{\hbar^2}{4m^2c^2}V'(r)\frac{\partial}{\partial r} + \frac{1}{2m^2c^2}\frac{1}{r}V'(r)\,\vec{L}.\vec{S} + V(r)\right]\psi_+ \ . \tag{20.244}$$

Le terme en $\vec{L}.\vec{S}$ est l'interaction spin-orbite cherchée, identique à l'expression (20.20) à condition de prendre strictement $g_{\rm e} = 2$ dans celle-ci.

Au même ordre d'approximation, le facteur de $\vec{p}^{\,2}$ peut et doit être simplifié : dans le terme $\frac{E_{\rm fr}-V}{2mc^2}\vec{p}^{\,2}\psi_+$, qui est visiblement d'ordre 1, on peut substituer l'expression de $\vec{p}^{\,2}\psi_+$ d'ordre zéro, égale à $2m(E_{\rm fr} - V)\psi_+$. Le terme qui en résulte fait apparaître $(E_{\rm fr} - V)^2\psi_+$, que l'on peut remplacer par $(\frac{\vec{p}^{\,2}}{2m})^2\psi_+$. Au total, (20.244) devient :

$$E_{\rm fr}\psi_+ = \left[\frac{\vec{p}^{\,2}}{2m} - \frac{\vec{p}^{\,4}}{8m^3c^2} + V(r) - \frac{\hbar^2}{4m^2c^2}\frac{\mathrm{d}V}{\mathrm{d}r}\frac{\partial}{\partial r} + \frac{1}{2m^2c^2}\frac{1}{r}\frac{\mathrm{d}V}{\mathrm{d}r}\vec{L}.\vec{S}\right]\psi_+ \ , \tag{20.245}$$

d'où par identification, le Hamiltonien faiblement relativiste :

$$\boxed{H_{\rm fr} \overset{\text{déf}}{=} \frac{\vec{p}^{\,2}}{2m} - \frac{\vec{p}^{\,4}}{8m^3c^2} + V(r) - \frac{\hbar^2}{4m^2c^2}\frac{\mathrm{d}V}{\mathrm{d}r}\frac{\partial}{\partial r} + \frac{1}{2m^2c^2}\frac{1}{r}\frac{\mathrm{d}V}{\mathrm{d}r}\vec{L}.\vec{S}} \tag{20.246}$$

Le terme en p^4 est une simple *correction de masse*, formellement identique à celui que donne le développement en $(v/c)^2$ de l'expression classique $(\gamma - 1)mc^2$, juste après le terme d'énergie cinétique.

Le quatrième terme est une correction relativiste à l'énergie potentielle, sans équivalent classique (il contient $\hbar$ en facteur). Il est usuellement donné sous une autre

forme, définissant *ipso facto* un Hamiltonien effectif en vue d'un traitement perturbatif des corrections faiblement relativistes. On peut en effet montrer (voir problème 20.8.5) que, dans une telle approximation, ce terme donne une correction proportionnelle à $|\psi_{nlm}(0)|^2$; de ce fait, elle est non-nulle exclusivement pour les états de type S ($l = 0$), les seuls donnant une densité de probabilité *finie* en $r = 0$. Cette correction peut donc s'écrire $\propto |\psi_{nlm}(\vec{r})|^2\,\delta(\vec{r})$, ou encore $\propto \psi_{nlm}(\vec{r})\,\delta(\vec{r})$. Par ailleurs, le potentiel du noyau $U(r) = |e|/(4\pi\varepsilon_0 r)$ satisfait l'équation de Poisson $\Delta U = -\rho(\vec{r})/\varepsilon_0 \equiv -|e|/\varepsilon_0)\delta(\vec{r})$, de sorte que l'énergie potentielle de l'électron $V(r)$ vérifie $\Delta V = 4\pi e'^2\delta(\vec{r})$; ceci permet d'exprimer la fonction de Dirac en terme de ΔV et d'écrire finalement le Hamiltonien *effectif* dans la limite faiblement relativiste :

$$\boxed{\tilde{H}_{\text{fr}} \stackrel{\text{déf}}{=} \frac{\vec{p}^2}{2m} + V(r) - \frac{\vec{p}^4}{8m^3c^2} + \frac{\hbar^2}{8m^2c^2}\Delta V + \frac{1}{2m^2c^2}\frac{1}{r}\frac{\mathrm{d}V}{\mathrm{d}r}\vec{L}.\vec{S}} \tag{20.247}$$

Ainsi transformé, ce terme porte le nom de *terme de Darwin* (voir problèmes 20.8.5 et 20.8.8).

20.8 Exercices et problèmes

20.8.1 Constantes du mouvement en théorie de Dirac

Trouver les commutateurs de $\vec{S}^2$ et $\vec{L}^2$ avec le Hamiltonien de Dirac H_{D}. Sans faire de calculs, mais en argumentant intuitivement et en prenant appui sur les résultats de Sommerfeld (Tome I, sous-section 7.3.3), expliquer pourquoi le vecteur de Lenz - Runge, $\vec{R}$ (voir (19.235)), n'est plus une constante du mouvement.

20.8.2 Homomorphisme SU(2) → SO(3)

Il s'agit de construire précisément la relation entre SO(3) et SU(2), et de montrer notamment qu'il existe une certaine application non-biunivoque entre ces deux groupes [12].

1. Soit $\mathcal{A}$ l'ensemble des matrices hermitiques 2×2 de trace nulle et $A \in \mathcal{A}$. Montrer que A peut s'écrire sous la forme :

$$A = \vec{a}.\vec{\sigma} \ , \tag{20.248}$$

où $\vec{\sigma}$ est l'ensemble des trois matrices de Pauli et $\vec{a}$ est un certain vecteur de $\mathbb{R}^3$; donner ses composantes a_u ($u = x,\ y,\ z$) en fonction des traces $\text{Tr}\,(A\sigma_u)$.

2. En déduire que $\mathcal{A}$ et $\mathbb{R}^3$ sont isomorphes.

3. Écrire la matrice A à l'aide des composantes de $\vec{a}$.

4. Exprimer le déterminant de A en fonction de $\|\vec{a}\|^2$.

5. Soit $U \in \mathrm{SU}(2)$ que l'on peut toujours écrire $U = \mathrm{e}^{\mathrm{i}\frac{\theta}{2}\vec{n}.\vec{\sigma}}$ où $\vec{n}$ est un vecteur unitaire de $\mathbb{R}^3$. À U, on associe par l'application $\mathcal{H}$ l'opération R_U, application dans $\mathcal{A}$ qui transforme A en A' suivant la définition :

$$U \in \mathrm{SU}(2) \xrightarrow{\mathcal{H}} R_U \;\; : A \xrightarrow{R_U} A' \equiv R_U(A) \stackrel{\text{déf}}{=} UAU^{-1} \;\; ; \tag{20.249}$$

montrer que $A' \in \mathcal{A}$. Comme $\mathcal{A}$ et $\mathbb{R}^3$ sont isomorphes, on garde la notation R_U pour désigner la transformation dans $\mathbb{R}^3$ qui transforme[103] $\vec{a}$ en $\vec{a}\,'$:

$$\vec{a} \xrightarrow{R_U} \vec{a}\,' = R_U(\vec{a}) \;\; . \tag{20.250}$$

6. Comparer $\|\vec{a}\,'\|^2$ et $\|\vec{a}\|^2$.

7. En raisonnant par continuité, montrer que Dét $R_U = +1$.

8. En déduire que $R_U \in \mathrm{SO}(3)$.

9. Montrer que $R_{U_1}R_{U_2} = R_{U_1U_2}$, ce qui établit que $\mathcal{H}$ est un homomorphisme.

10. Quelle est la relation entre R_U et R_{-U} ?

11. L'application $\mathcal{H}$ est-elle un isomorphisme ?

12. On note $R_{U,\,uv}$ la matrice de R_U telle que les composantes de $\vec{a}$ et $\vec{a}\,'$ satisfont :

$$a'_u = \sum_v R_{U,\,uv}a_v \;\; . \tag{20.251}$$

   (a) Exprimer $R_{U,\,uv}$ sous la forme d'une trace.

   (b) Préciser ces résultats en écrivant la matrice R_U dans le cas d'une rotation d'angle θ autour de Oz.

   (c) Comment se compare cette dernière matrice avec $\mathcal{R}_{\mathrm{O}z,\,\theta}$ pour un moment cinétique $\vec{J}$ quand $J = 1$?

20.8.3 Harmoniques sphériques spinorielles

Soit $\vec{J} = \vec{L} + \vec{S}$, avec $S = 1/2$. Il s'agit de trouver les vecteurs propres $|jm\rangle$ de $\vec{J}^2$, J_z à partir des états $Y_{l m_l}(\theta,\,\phi) \otimes |m_s = \pm\frac{1}{2}\rangle$ (voir éq. (20.193)).

1. Comment doit-on combiner m_l et m_s pour avoir la valeur propre $m\hbar$ pour J_z ?

2. En déduire que la diagonalisation de $\vec{J}^2$ s'effectue dans un espace de dimension égale à 2 (en préciser une base).

[103]on note $\vec{a}\,'$ le vecteur tel que $A' = \vec{a}\,'.\vec{\sigma}$.

3. En utilisant :
$$\vec{J}^2 = \vec{L}^2 + \vec{S}^2 + 2L_zS_z + L_+S_- + L_-S_+ \tag{20.252}$$
trouver la matrice de $\vec{J}^2$ sur la base déterminée précédemment.

4. Écrire l'équation aux valeurs propres pour $\vec{J}^2$ et retrouver que les deux seules valeurs possibles de j sont $l \pm \frac{1}{2}$.

5. Trouver effectivement les vecteurs propres, et comparer aux expressions (20.193).

20.8.4 Limite faiblement relativiste de la densité et du courant

En considérant le cas d'une particule libre, trouver la limite faiblement relativiste de la densité ρ et du courant $\vec{j}$. Comparer aux expressions connues dans le cas non-relativiste (voir Tome I, éq. (9.92)).

20.8.5 Correction de Darwin

Il s'agit de détailler le passage du Hamiltonien donné en (20.246) à l'opérateur effectif défini juste après, (20.247). La méthode consiste à trouver les corrections d'énergie en les identifiant à leur valeur moyenne dans les états non-relativistes. Comme on le verra au chapitre 23, ceci est la règle de base, s'agissant d'un niveau non-dégénéré, pour trouver les corrections au premier ordre en théorie des perturbations stationnaires.

En désignant par $\langle ... \rangle$ une telle valeur moyenne, l'énergie d'un certain état non dégénéré prend la forme :
$$E_{\text{fr}} = E_{c=\infty} - \frac{1}{8m^3c^2}\langle \vec{p}^4 \rangle - \frac{\hbar^2}{4m^2c^2}\langle \frac{\mathrm{d}V}{\mathrm{d}r}\frac{\partial}{\partial r}\rangle + ... \tag{20.253}$$
où les ... désignent la correction de spin-orbite. Dans la suite, on ne considère que le terme $\propto \langle \frac{\mathrm{d}V}{\mathrm{d}r}\frac{\partial}{\partial r}\rangle$, dont il s'agit de montrer qu'il agit comme le terme de Darwin :
$$V_{\text{Darwin}} = \frac{\hbar^2}{8m^2c^2}\Delta V(r) \tag{20.254}$$
où Δ désigne le Laplacien et $V(r)$ l'énergie potentielle électrostatique de l'électron dans le champ du noyau (proton).

1. Montrer que $rV'' + 2V' = 0$.

2. Écrire l'expression formelle de la valeur moyenne $\langle \frac{\mathrm{d}V}{\mathrm{d}r}\frac{\partial}{\partial r}\rangle$ dans un état d'ordre zéro (strictement non-relativiste) $\psi_{nlm}(\vec{r}) = R_{nl}(r)Y_{lm}(\theta, \phi)$.

3. En effectuant des intégrations par parties, montrer que cette moyenne est égale à :
$$\langle \frac{\mathrm{d}V}{\mathrm{d}r}\frac{\partial}{\partial r}\rangle = -\frac{1}{2}e'^2\,[R_{nl}(0)]^2 \tag{20.255}$$

4. Écrire l'équation de Poisson satisfaite par le potentiel électrostatique du noyau, $U(r)$, et en déduire l'équation pour l'énergie potentielle $V(r)$.

5. En déduire que la correction coïncide avec le terme effectif de Darwin défini en (20.254), et que, en vue d'un tel traitement perturbatif, les deux Hamiltoniens (20.246) et (20.247) sont équivalents.

20.8.6 Ordres de grandeur des corrections relativistes

En l'absence de champ magnétique, la limite faiblement relativiste de l'équation de Dirac produit le Hamiltonien suivant pour l'hydrogène[104] :

$$H_{\text{fr}} = H_0 - \frac{p^4}{8m^3c^2} + a(r)\vec{L}.\vec{S} + \frac{\hbar^2}{8m^2c^2}\Delta V(r) \ , \tag{20.256}$$

où :

$$H_0 = \frac{\vec{p}^{\,2}}{2m} + V(r) \ , \qquad V(r) = -\frac{e'^2}{r} \ , \qquad a(r) = \frac{1}{2m^2c^2}\,\frac{1}{r}\,\frac{\mathrm{d}V}{\mathrm{d}r} \ , \tag{20.257}$$

et où Δ désigne le Laplacien dans $\mathbb{R}^3$. Dans la suite, on désigne par E_n les énergies des états $|nLM_LSM_S\rangle$ de la théorie non relativiste :

$$E_n = -\frac{me'^4}{2n^2\hbar^2} = -\frac{\alpha^2}{2n^2}mc^2 \simeq -\frac{13,6}{n^2}\ \text{eV} \ . \tag{20.258}$$

α désigne la constante de structure fine ($\alpha = \frac{e'^2}{\hbar c} \simeq \frac{1}{137}$)

Dans l'expression (20.256), le terme en p^4 vient de la "variation de la masse" ; le dernier terme, appelé *terme de contact de Darwin*, peut être interprété comme une non-localité de l'interaction noyau - électron, impliquant un petit volume de l'ordre de λ_{C}^3, où λ_{C} est la longueur d'onde Compton de l'électron. Enfin, le terme en $\vec{L}.\vec{S}$ représente l'interaction spin-orbite, dont on sait qu'elle donne des corrections d'ordre α^2 relativement aux énergies obtenues dans la limite $c = +\infty$.

1. Par le théorème du Viriel, l'énergie cinétique de l'électron est du même ordre de grandeur que E_n. En déduire que :

$$\frac{p^4}{8m^3c^2} \sim \alpha^2 E_n \ . \tag{20.259}$$

2. Sachant que le Laplacien de r^{-1} s'exprime comme suit à l'aide de la fonction de Dirac :

$$\Delta\,\frac{1}{r} = -4\pi\,\delta(\vec{r}) \ , \tag{20.260}$$

trouver l'expression de la valeur moyenne du terme de Darwin dans un état atomique $\psi(\vec{r})$

[104] à la constante additive mc^2 près.

3. En s'appuyant sur le fait que l'électron, dans un état lié pas trop excité, est essentiellement localisé dans un volume d'ordre a_0^3, quel est l'ordre de grandeur de $|\psi(\vec{0})|^2$ pour un état S ($L = 0$) ?

4. En déduire l'ordre de grandeur du terme de Darwin. Combien vaut ce terme pour un état de moment cinétique orbital non-nul ($L > 0$) ?

5. Trouver toutes les corrections des états $n = 1$ et $n = 2$; identifier le doublet H_α dont les deux composantes sont séparées de $0,365\,\mathrm{cm}^{-1} \simeq 10\,950\,\mathrm{MHz}$ (voir Tome I, p. 221).

20.8.7 Mesure de l'anomalie magnétique de l'électron

Le moment magnétique de l'électron (charge e, masse m), $\vec{\mu}$, est relié à son spin $\vec{S}$ par :

$$\vec{\mu} = g_\mathrm{e}\frac{e}{2m}\,\vec{S} \ . \tag{20.261}$$

Alors que la théorie de Dirac donne $g_\mathrm{e} = 2$, l'expérience montre que $g_\mathrm{e} \gtrsim 2$; pour cette raison, on pose $g_\mathrm{e} = 2(1+a)$, de sorte que :

$$\vec{\mu} = (1+a)\frac{e}{m}\,\vec{S} \equiv \gamma\,\vec{S} \ , \tag{20.262}$$

où γ est le facteur gyromagnétique de l'électron. a est petit devant l'unité, de l'ordre de la constante de structure fine α ; plus précisément, on sait que :

$$a = \frac{\alpha}{2\pi} + \mathcal{O}(\alpha^2) \ , \qquad \frac{\alpha}{2\pi} \simeq 1,16 \times 10^{-3} \ . \tag{20.263}$$

Il s'agit de voir comment l'étude de la dynamique de l'électron en présence d'un champ magnétique uniforme et statique $\vec{\mathcal{B}}$ parallèle à l'axe Oz permet de déterminer l'anomalie a. On introduira la pulsation-cyclotron ω définie comme :

$$\omega = \frac{e\mathcal{B}}{m} \ . \tag{20.264}$$

1. (a) Comment s'exprime la vitesse $\vec{v}$ de l'électron en fonction de $\vec{p}$ et $\vec{A}$?

   (b) Le Hamiltonien de l'électron est :

   $$H = \frac{1}{2m}(\vec{p} - e\vec{A})^2 - \gamma\mathcal{B}S_z \ . \tag{20.265}$$

   Interpréter les différents termes.

   (c) $\vec{r}$ désignant le rayon-vecteur de l'électron, on choisit le potentiel-vecteur comme suit :

   $$\vec{A} = \frac{1}{2}\,\vec{\mathcal{B}} \times \vec{r} \ . \tag{20.266}$$

   Écrire les composantes de $\vec{A}$.

(d) Calculer les commutateurs des composantes de la vitesse entre elles, ainsi que le commutateur $[\vec{v}, \vec{v}^{\,2}]$.

(e) En déduire l'équation de Heisenberg[105] pour les trois composantes de $\vec{v}$. Montrer que l'équation du mouvement de la vitesse $\vec{v}$ peut se mettre sous la forme :

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \blacksquare\,(\vec{\bullet} \times \vec{\mathcal{B}}) \ . \tag{20.267}$$

Que sont $\blacksquare$ et $\vec{\bullet}$? Décrire le mouvement de $\vec{v}$.

(f) Reprendre les points de la question précédente pour le spin $\vec{S}$.

(g) Que serait le mouvement conjoint de $\vec{v}$ et $\vec{S}$ si a était nul ?

2. Les résultats précédents conduisent à s'intéresser à la quantité $\vec{S}.\vec{v}$, qui serait constante dans le temps si a était nul. Quelle est, avec $a \neq 0$, l'équation de Heisenberg pour $\vec{S}.\vec{v}$?

3. Cette dernière équation fait apparaître le produit vectoriel $\vec{S} \times \vec{v}$. Un calcul simple montre que :

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(\vec{S} \times \vec{v}) = \frac{e}{m}\,[a(\vec{S}.\vec{v})\,\vec{\mathcal{B}} - (1+a)(\vec{\mathcal{B}}.\vec{v})\,\vec{S} + (\vec{\mathcal{B}}.\vec{S})\,\vec{v}] \ . \tag{20.268}$$

En déduire une équation différentielle du second ordre pour $\vec{S}.\vec{v}$.

4. En remarquant que $S_z v_z$ est une constante du mouvement, en déduire que, pour un état initial donné, la valeur moyenne de $\vec{S}.\vec{v}$ est donnée par :

$$\langle \vec{S}.\vec{v} \rangle(t) = A + C\,\cos(a\omega t + \phi) \ , \tag{20.269}$$

où A et C sont des constantes ici sans intérêt.

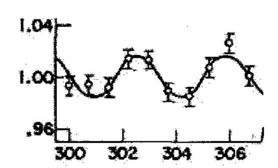


Figure 20.3: Oscillation de $\langle \vec{S}.\vec{v} \rangle$ en fonction du temps en μs (figure extraite de l'article de Wilkinson et Crane [91]).

5. La mesure de la quantité $\langle \vec{S}.\vec{v} \rangle$ montre qu'elle oscille à la période $T \simeq 3\,\mu$s (voir la figure 20.3). Calculer numériquement a pour $\mathcal{B} = 0,0094$ T et comparer avec la prévision théorique exprimée par (20.263).

[105] Pour simplifier les notations, on pourra omettre l'indice $_\mathrm{H}$ dans la notation des observables dans l'image (ou point de vue) de Heisenberg.

20.8.8 Transformation de Foldy – Wouthuysen

La transformation de Foldy - Wouthuysen [92] est une méthode systématique pour, partant de l'équation de Dirac, engendrer les corrections relativistes sous forme d'un développement en puissances (entières) du petit paramètre v/c. Il s'agit ici simplement de comprendre le principe de cette approche, et aussi de se familiariser un peu avec l'algèbre des matrices $\vec{\alpha}$ et β de Dirac ; pour plus de détails, on peut consulter [15], section XX-V.

Partant de $H_\mathrm{D}|\Psi\rangle = E|\Psi\rangle$, où H_D est défini en (20.147), la méthode consiste à poser $|\Psi'\rangle \stackrel{\text{déf}}{=} U|\Psi\rangle$, où U est une certaine transformation, pour écrire l'équation aux vecteurs propres sous la forme :

$$H'_\mathrm{D}|\Psi'\rangle = E|\Psi'\rangle \ , \qquad H'_\mathrm{D} \stackrel{\text{déf}}{=} U H_\mathrm{D} U^{-1} \ . \tag{20.270}$$

1. On choisit :
$$U \stackrel{\text{déf}}{=} \mathrm{e}^{\Omega} \ , \quad \Omega \stackrel{\text{déf}}{=} \frac{\beta\vec{\alpha}.\vec{p}}{2mc} \ , \tag{20.271}$$
où $\vec{\alpha}$ et β sont les opérateurs de Dirac, et où $\vec{p}$ est le moment conjugué. Montrer que U est une transformation unitaire.

2. Montrer que U s'écrit :
$$U = u_0(p)\,\mathbf{1}_4 + u(p)\,\beta\vec{\alpha}.\vec{p} \ , \tag{20.272}$$
où $u_0(p)$ et $u(p)$ sont des fonctions à déterminer ($p \stackrel{\text{déf}}{=} \|\vec{p}\|$). Utiliser cette forme pour vérifier que $UU^\dagger = U^\dagger U = \mathbf{1}_4$.

3. Montrer que si B et A anticommutent, on a $B\mathrm{e}^{-A} = \mathrm{e}^{A}B$. Dans ces conditions, combien vaut $\mathrm{e}^{A}B\mathrm{e}^{-A}$?

4. Montrer que $\beta\vec{\alpha}.\vec{p}$ anticommute avec $\beta mc^2 + c\vec{\alpha}.\vec{p}$.

5. En déduire que $(\beta mc^2 + c\vec{\alpha}.\vec{p})' = \mathrm{e}^{2\Omega}(\beta mc^2 + c\vec{\alpha}.\vec{p})$

6. Écrire $(\beta mc^2 + c\vec{\alpha}.\vec{p})'$ en forme linéaire de β et $\vec{\alpha}.\vec{p}$.

7. En utilisant le développement (13.204) (Tome I, p. 468), obtenir les trois premiers termes de $V' \stackrel{\text{déf}}{=} UVU^\dagger$.

8. En déduire, à cet ordre, le Hamiltonien transformé H'_D, et y reconnaître le terme de Darwin.

20.8.9 Zitterbewegung

Le terme *Zitterbewegung* désigne un *mouvement de tremblement* ; l'effet est une conséquence immédiate de l'existence des états d'énergie négative de la théorie de Dirac. Interprété historiquement par Schrödinger comme une interférence entre les deux classes

d'états d'énergies positives et négatives, cet effet doit être considéré comme un *artefact* sans signification physique, et comme un symptôme de plus des difficultés insurmontables de la théorie tant qu'elle est interprétée comme une théorie à une particule.

Dans la suite, on raisonne avec une particule libre, décrite par l'équation de Dirac :

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\Psi(\vec{r},\, t)\rangle = H_{\mathrm{D}}|\Psi(\vec{r},\, t)\rangle \ , \tag{20.273}$$

où H_{D} est défini en (20.146). Le cas échéant, on se placera dans la représentation standard où les opérateurs $\vec{\alpha}$ et β sont représentés par les matrices données en (20.145).

1. Soit A une observable quelconque, et soit $A(t)$ sa représentation dans le point de vue de Heisenberg (voir Tome I, sous-section 14.1.2) ; écrire l'équation de Heisenberg pour $A(t)$.

2. En particulier, montrer que la représentation de Heisenberg de la position, $\vec{r}(t)$, obéit à :
$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{r}(t) = c\vec{\alpha} \ . \tag{20.274}$$

3. Écrire de même l'équation pour $\vec{\alpha}(t)$, et la mettre sous la forme :
$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{\alpha}(t) = \frac{2\mathrm{i}}{\hbar}(c\vec{p} - \vec{\alpha}H_{\mathrm{D}}) \ . \tag{20.275}$$

4. Intégrer cette équation en posant $\vec{\alpha}(t) = cH_{\mathrm{D}}^{-1}\vec{p} + \vec{a}(t)$.

5. En déduire l'expression de $\vec{r}(t)$:
$$\vec{r}(t) = \vec{r} + \frac{c^2\vec{p}}{H_{\mathrm{D}}}t + \frac{\mathrm{i}\hbar c}{2}\left(\frac{c\vec{p}}{H_{\mathrm{D}}} - \vec{\alpha}\right)\frac{1}{H_{\mathrm{D}}}\left(1 - \mathrm{e}^{-\frac{2\mathrm{i}}{\hbar}H_{\mathrm{D}}t}\right) \ , \tag{20.276}$$
qui est aussi :
$$\vec{r}(t) = \vec{r} + \frac{c^2\vec{p}}{H_{\mathrm{D}}}t + \frac{\mathrm{i}\hbar c}{2}\left[\vec{\alpha}(t) - \vec{\alpha}\right]\frac{1}{H_{\mathrm{D}}} \ . \tag{20.277}$$

6. Dans toute la suite, on décompose l'état initial à quatre composantes, $|\Psi_0(\vec{r})\rangle$, sous la forme $C_+|\Psi_+(\vec{r})\rangle + C_-|\Psi_-(\vec{r})\rangle$, où les $|\Psi_\pm(\vec{r})\rangle$ sont des vecteurs normalisés appartenant respectivement aux deux sous-espaces d'énergies positives et négatives, représentant chacun un paquet d'ondes construit sur les états propres de H_{D}.
   (a) Expliquer pourquoi $\langle\Psi_\pm|\vec{p}|\Psi_\mp\rangle = 0$.
   (b) On pose $\vec{r}_0 \stackrel{\text{déf}}{=} \langle\Psi_0|\vec{r}|\Psi_0\rangle$, $\vec{\varpi} \stackrel{\text{déf}}{=} \langle\Psi_0|\frac{c\vec{p}}{H_{\mathrm{D}}}|\Psi_0\rangle$ et :
$$\vec{d}(t) \stackrel{\text{déf}}{=} \frac{\mathrm{i}\hbar c}{2}\sum_{\varepsilon} C_\varepsilon^* C_{-\varepsilon}\langle\Psi_\varepsilon|[\vec{\alpha}(t) - \vec{\alpha}]\frac{1}{H_{\mathrm{D}}}|\Psi_{-\varepsilon}\rangle \tag{20.278}$$

Montrer que la valeur moyenne de la coordonnée est donnée par :

$$\langle \vec{r} \rangle(t) = \vec{r}_0 + \vec{\varpi}\, ct + \vec{d}(t) \ . \tag{20.279}$$

7. Comment se simplifie l'expression précédente dans le cas où l'état initial ne contient que des états d'énergie positive (resp. négative), soit si $C_- = 0$ (resp. $C_+ = 0$) ? Analyser sommairement ce résultat.

8. On suppose au contraire que $|C_+ C_-| \neq 0$; expliquer pourquoi la position moyenne effectue des oscillations très rapides autour d'une trajectoire moyenne semblable à celle obtenue dans la question précédente – c'est le *Zitterbewegung*, conséquence directe de l'interférence entre les états d'énergie positive et négative.

9. Dans le cas où les états $|\Psi_\pm\rangle$ ont des énergies (moyennes) voisines de $\pm mc^2$, justifier qualitativement que l'on puisse alors écrire :

$$\|\vec{d}(t)\| \sim \lambda_{\mathrm{C}} \sin(2\hbar^{-1} mc^2 t) \ , \tag{20.280}$$

où λ_{C} est la longueur d'onde Compton de la particule ; expliquer pourquoi l'amplitude de ces oscillations très rapides est très petite devant la largeur Δr du paquet d'ondes. Estimer la fréquence ν des oscillations de *Zitterbewegung*.

20.8.10 Puits carré en théorie de Dirac

Soit à résoudre l'équation de Dirac pour une particule de masse m dans un puits de potentiel sphérique de profondeur $-V_0$ ($V_0 > 0$) et de rayon a. Afin d'éviter les difficultés liées aux états d'énergie négative, si problématiques, on s'intéresse uniquement aux états propres (liés) d'énergie E comprise entre mc^2 et $mc^2 - V_0$, supposant de surcroît que la profondeur du puits est plutôt petite devant mc^2. On pose :

$$E = mc^2 \cos\phi \ , \qquad V(r) = mc^2 v(r) \ , \qquad V_0 = mc^2\, v_0 \ , \qquad \phi \in \mathbb{R} \ . \tag{20.281}$$

On note $F(r)$ et $G(r)$ les fonctions radiales à déterminer (voir section 20.6, et notamment (20.213)). Des rappels utiles concernant les fonctions de Bessel sont donnés en fin d'énoncé, qu'il convient de consulter au fil de l'élaboration de la solution.

Par ailleurs, $\mathsf{q} \in \mathbb{Z}^*$ note toujours la valeur propre de l'opérateur Q défini en (20.206) ; il est rappelé que $\mathsf{q} = \pm(j + \frac{1}{2})$, où j désigne ici un demi-entier positif, associé à la valeur propre de $\vec{J}^2$. On pourra poser :

$$\lambda \stackrel{\text{déf}}{=} \frac{\hbar}{mc} \ . \tag{20.282}$$

À quelle grandeur physique connue la quantité λ est-elle reliée ?

1. En utilisant le fait que $V(r)$ est constant par morceaux, éliminer la fonction $G(r)$ pour obtenir une équation différentielle du second ordre n'impliquant que la fonction $F(r)$.

2. Écrire explicitement cette équation dans les deux régions $r < a$ et $r > a$.

3. *Région intérieure*

   (a) En posant $\rho = k_0 r$, $F(r) = \bar{F}(\rho)$, montrer que l'équation pour $F(r)$ se ramène à l'équation sphérique de Bessel. Identifier le paramètre k_0 en fonction de v_0 et de ϕ ; montrer qu'il est réel, et peut donc être pris positif dans la suite.

   (b) Utiliser la condition aux limites en $r = 0$ pour montrer que la solution $F(r)$ est de la forme $Aj_{l_\mathsf{q}}(\rho)$ où $l_\mathsf{q} \in \mathbb{N}$ dépend du signe de q.

   (c) En déduire la fonction $G(r)$.

   (d) Identifier les grandes et les petites composantes.

4. *Région extérieure*

   (a) Montrer que l'équation pour $F(r)$ se ramène à l'équation sphérique de Bessel en posant cette fois $\rho = \mathrm{i}kr$, $F(r) = \bar{F}(\rho)$, avec $k > 0$.

   (b) En déduire que la solution $F(r)$ implique essentiellement la fonction de Bessel modifiée $K_\nu(kr)$, où l'indice ν dépend du signe de q.

   (c) En déduire la fonction $G(r)$ pour $r > a$.

5. Écrire la condition de raccordement en $r = a$.

6. Expliciter cette condition dans le cas des états S ($l = 0$), soit $\mathsf{q} = 1$.

7. Expliquer comment elle permet de trouver les valeurs de l'énergie.

8. Trouver la valeur minimum de V_0 qui permet de lier une particule de masse m, et comparer au traitement non-relativiste (voir (19.106)).

9. Expliquer comment, partant de ces résultats, on peut très rapidement obtenir les états sphériques d'une particule libre dans $\mathbb{R}^3$, équivalents relativistes des états décrits dans la section 19.1.3.

Rappels sur les fonctions de Bessel

On a déjà rencontré ici et là les fonctions de Bessel (voir notamment chapitre 19). Les éléments suivants sont utiles pour le problème, certains étant répétés pour la seule commodité, ce qui ne dispense pas de consulter un ouvrage traitant plus particulièrement de ces fonctions et de leurs propriétés.

1. L'équation *sphérique* de Bessel est :

$$f''(z) + \frac{2}{z}f'(z) + [1 - \frac{l(l+1)}{z^2}]f(z) = 0 \qquad (l \in \mathbb{N}) \ . \tag{20.283}$$

   Elle possède deux solutions linéairement indépendantes, le plus souvent notées $j_l(z)$ et $y_l(z)$, qui sont reliées aux fonctions de Bessel $J_\nu(z)$:

$$j_l(z) \stackrel{\text{déf}}{=} \sqrt{\frac{\pi}{2z}}\, J_{l+\frac{1}{2}}(z) \ , \qquad y_l(z) \stackrel{\text{déf}}{=} (-1)^{l+1}\sqrt{\frac{\pi}{2z}}\, J_{-l-\frac{1}{2}}(z) \ . \tag{20.284}$$

2. La fonction $J_\nu(z)$ admet le développement en série :

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{n\in\mathbb{N}} \frac{(-1)^n}{n!\Gamma(n+1+\nu)} \left(\frac{z}{2}\right)^{2n} \quad , \tag{20.285}$$

où Γ est la fonction d'Euler de deuxième espèce.

3. Les fonctions $j_l(z)$ et $y_l(z)$ peuvent s'obtenir à partir des relations suivantes :

$$j_l(z) = (-z)^n \left(\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^n \frac{\sin z}{z} \quad , \qquad y_l(z) = -(-z)^n \left(\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^n \frac{\cos z}{z} \quad . \tag{20.286}$$

4. Les fonctions de Bessel *modifiées* apparaissent naturellement quand partant d'un argument réel x, on est contraint d'introduire des fonctions du genre $J_\nu(\mathrm{i}x)$; on définit deux classes de fonctions de Bessel modifiées :

$$I_\nu(x) = \mathrm{i}^{-\nu}\, J_\nu(\mathrm{i}x) \quad , \qquad K_\nu(x) = \frac{\pi}{2\sin\nu\pi}[I_{-\nu}(x) - I_\nu(x)] \quad . \tag{20.287}$$

Ces fonctions ont pour comportement à l'infini :

$$I_\nu(x) \sim \frac{1}{\sqrt{2\pi x}}\,\mathrm{e}^x \quad , \qquad K_\nu(x) \sim \sqrt{\frac{\pi}{2x}}\,\mathrm{e}^{-x} \qquad (x \to +\infty) \quad . \tag{20.288}$$

Quand $\nu = n + \frac{1}{2}$ ($n \in \mathbb{N}$), la fonction K_ν a le développement :

$$K_{n+\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}}\,\mathrm{e}^{-x} \sum_{k=0}^{n} \frac{(n+k)!}{k!(n-k)!}(2z)^{-k} \quad . \tag{20.289}$$

5. $Z_\nu(z)$ désignant une fonction de Bessel quelconque, elle satisfait (notamment) la relation suivante :

$$zZ'_\nu(z) + \nu Z_\nu(z) = zZ_{\nu-1}(z) \quad . \tag{20.290}$$

20.8.11 • Paquet d'ondes gaussien de Dirac

Partant des solutions en ondes planes pour une particule libre, on forme un paquet d'ondes avec les solutions d'énergie positive, paramétrée suivant $E = mc^2\cosh\phi$, et d'hélicité positive $|\psi^{(\vec{k})}_{+,+1}(\vec{r})\rangle$, dont l'expression est donnée en (20.181) :

$$|\Psi(\vec{r})\rangle = \int_{\mathbb{R}} A(\vec{k})|\psi^{(\vec{k})}_{+,+1}(\vec{r})\rangle\,\mathrm{d}k \quad . \tag{20.291}$$

Dans toute la suite, on ne considère que des vecteurs $\vec{k}$ parallèles à la constante du mouvement $\vec{p}$, de module $p = mc\sinh\phi$ et définisant l'axe Oz : $\vec{k} = \frac{k}{p}\vec{p}$, $k \in \mathbb{R}$; dès lors, on considère le paquet d'ondes unidimensionnel :

$$|\Psi(z)\rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} A(k)\frac{\mathrm{e}^{\mathrm{i}kz}}{\sqrt{\cosh\phi}} \begin{bmatrix} \cosh\frac{\phi}{2} \\ 0 \\ \sinh\frac{\phi}{2} \\ 0 \end{bmatrix} \mathrm{d}k \quad . \tag{20.292}$$

On choisit une enveloppe $A(k)$ gaussienne :

$$A(k) = A_0 \,\mathrm{e}^{-\frac{(k-k_0)^2}{4\Delta k^2}} \quad . \tag{20.293}$$

1. Normaliser $|\Psi(\vec{r})\rangle$ en trouvant la constante A_0.
2. La valeur moyenne d'une observable Ω, $\langle\Omega\rangle$, est égale à $\int_{\mathbb{R}} \langle\Psi(z)|\Omega|\Pi(z)\rangle \,\mathrm{d}z$; écrire l'expression formelle de cette moyenne.
3. Montrer que la valeur moyenne de la coordonnée est :

$$\langle z\rangle = \mathrm{i}\int_{\mathbb{R}} A^*(k)A'(k)\,\mathrm{d}k \tag{20.294}$$

   (on utilisera la relation $\int_{\mathbb{R}} z\mathrm{e}^{\mathrm{i}kz}\mathrm{d}z = -2\mathrm{i}\pi\delta'(k)$). Combien vaut $\langle z\rangle$ avec l'enveloppe gaussienne ?
4. Établir l'égalité :

$$\langle z^2\rangle = \int_{\mathbb{R}} |A'(k)|^2\,\mathrm{d}k \quad . \tag{20.295}$$

5. Trouver les valeurs moyennes dans l'état $|\Psi(\vec{r})\rangle$:
   (a) de la matrice α_z (l'exprimer en fonction de ϕ) ; en déduire la moyenne de la composante suivant Oz de la vitesse $\vec{v}$;
   (b) de l'impulsion p ;
   (c) du carré de l'énergie, $\langle E^2\rangle$.
6. En déduire l'écart quadratique de l'impulsion, Δp^2.
7. Quel est l'écart quadratique Δz^2 de la coordonnée le long de Oz ?
8. Combien vaut le produit $\Delta z\,\Delta p$?
9. Quelle est l'expression formelle de l'état à l'instant t, $|\Psi(z,\,t)\rangle$, quand $|\Psi(z)\rangle$ est choisi comme état initial ? Sans faire de calcul, expliquer pourquoi, dans la limite non-relativiste, $|\Psi(z,\,t)\rangle$ redonne le paquet d'ondes gaussien d'une particule de spin $+\hbar/2$ se déplaçant à la vitesse $\hbar k_0/m$ et dont la largeur caractéristique est $\Delta z(t) = \sqrt{\Delta z^2 + (\Delta p/m)^2 t^2}$.
10. Toujours sans faire de calcul, confirmer ce dernier résultat en analysant l'expression de la densité $\rho(x,\,t)$ dans la limite $c \to +\infty$.

physique

Chapitre 21

Illustration des postulats de la Mécanique quantique

Il s'agit d'abord de présenter quelques illustrations simples des postulats en décrivant des expériences récentes. Deux des sujets abordés (effet Zénon et cryptographie) discutent d'idées introduites ces dernières décennies. En revanche, les sauts quantiques reposent sur un concept énoncé par Bohr (1913), mais n'ont pu être observés que dans le passé proche. Ensuite, les bases de la théorie de la décohérence sont expliquées, avant une présentation de la notion d'intrication et un retour sur le paradoxe EPR, les inégalités de Bell et l'apport des expériences d'Aspect.

21.1 L'effet Zénon quantique

L'effet Zénon[1] quantique a été précisément formulé en 1977 [94], mais certains travaux antérieurs [95] avaient déjà évoqué ce qui paraît à première vue constituer un paradoxe : un système quantique soumis à une succession de mesures très rapprochées n'évolue pas dans le temps[2]. L'idée est facile à comprendre. Toute opération de mesure projette dans l'état propre normalisé de l'observable mesurée, associé à la valeur propre obtenue en tant que résultat de la mesure effectuée. La vitesse d'évolution du vecteur d'état étant finie, la probabilité d'obtenir la même valeur lors d'une deuxième mesure rapprochée dans le

[1]Le paradoxe de Zénon *classique* repose sur la négation *a priori* de la possibilité de diviser à l'infini l'espace et le temps, négation qui débouche sur l'impossibilité pour Achille de rattraper la tortue. Il en résulte des affirmations paradoxales tendant toutes à montrer que le mouvement est impossible (voir par exemple [93]). Il est intéressant de savoir que Schrödinger revint dans les années 30 sur la question de représenter l'espace et le temps par des grandeurs *continues*.

[2] *"It is common knowledge that a watched kettle never boils."* [96].

temps est relativement grande – en quelque sorte, la deuxième mesure *freine* l'évolution. Si on procède à une série de mesures très voisines en temps, on conçoit que se produise un ralentissement de la dynamique. L'argument qui suit montre, dans un cas d'école, que dans la limite d'un nombre infini de mesures effectuées entre deux instants fixés t_i et t_f, la probabilité pour le système d'être à l'instant t_f dans le *même* état qu'à l'instant t_i ... est égale à 1.

De façon plus précise, cet effet paradoxal peut se formuler comme suit. Soit un système quantique, de Hamiltonien H, initialement ($t_\mathrm{i} = 0$) dans l'état $|\psi_0\rangle$, état supposé instable pour fixer les idées[3]. Le projecteur $P_0 \stackrel{\text{déf}}{=} |\psi_0\rangle\,\langle\psi_0|$ est un opérateur hermitique, dont les valeurs propres ϖ_λ sont 0 et 1 :

$$\varpi_1 = 1 : \quad P_0\,|\psi_0\rangle = |\psi_0\rangle \ ; \qquad \varpi_0 = 0 : \quad P_0\,|\phi\rangle = 0 \quad \forall\,|\phi\rangle \perp |\psi_0\rangle \ . \tag{21.1}$$

La valeur propre 0 est en général fortement dégénérée (sauf si l'espace des états $\mathcal{E}$ est de dimension égale à 2). En tout cas, P_0 est un opérateur dont toutes les valeurs propres sont réelles et dont l'ensemble des vecteurs propres constitue une base complète de $\mathcal{E}$: c'est donc une observable au sens de la Mécanique quantique. De ce fait, il est licite d'appliquer formellement les postulats concernant les résultats d'une opération de mesure de P_0 effectuée sur un système dans l'état $|\Psi\rangle$ donné :

- le résultat d'une mesure est l'une des valeurs propres ϖ_λ de P_0.
- La probabilité d'obtenir le résultat ϖ_λ lors d'une mesure est le module au carré de la composante de $|\Psi\rangle$ sur le vecteur propre correspondant, $|\varpi_\lambda\rangle$.
- Juste après la mesure dont le résultat est $\varpi_1 = 1$, le système se trouve dans l'état $|\varpi_\lambda = 1\rangle \equiv |\psi_0\rangle$ (postulat de réduction du paquet d'ondes en l'absence de dégénérescence) ; si le résultat est $\varpi_0 = 0$, le système est, juste après cette mesure, dans l'état (normalisé) qui est la projection de l'état juste avant la mesure sur le sous-espace $\{|\varpi_\lambda = 0\rangle\}$ (réduction du paquet d'ondes en présence de dégénérescence).

On sait que si cette dernière affirmation a de bien étranges conséquences (le *Chat de Schrödinger !*), elle est toutefois une nécessité logique pour que la théorie quantique ne soit pas absurde, physiquement parlant. En effet, si la mesure à l'instant t_0 d'une certaine observable A sur le système dans l'état $|\Psi\rangle$ a produit la valeur non-dégénérée a_0, une deuxième mesure effectuée immédiatement après doit évidemment redonner la *même* valeur (avec probabilité 1). Ceci n'est possible que si à l'issue de la première mesure le système *est* dans l'état $|a_0\rangle$, propre de A avec la valeur propre a_0 ; la première mesure a donc bien effectué la projection :

$$|\Psi(t_0)\rangle \ \longrightarrow \ |a_0\rangle \ \equiv \ |\Psi(t_0 + 0)\rangle \ . \tag{21.2}$$

Effectuons une mesure unique de P_0 à l'instant t, le système étant parti à $t_\mathrm{i} = 0$ de l'état $|\psi_0\rangle$; le descendant de l'état de départ est $U(t)|\psi_0\rangle$ où $U(t) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Ht}$. Le résultat

[3] Cette hypothèse n'est pas essentielle, mais elle en rajoute au spectaculaire : une particule instable que l'on observe en continu est stable !

de cette mesure est donc :

$$1 \text{ avec la probabilité } p_1 = |\langle \varpi_1 | U(t) | \psi_0 \rangle|^2 \tag{21.3}$$

$$0 \text{ avec la probabilité } p_0 = 1 - p_1 \ . \tag{21.4}$$

Quand on a trouvé 1, on convient de dire que le résultat de la mesure est *positif.* $|\psi_0\rangle$ étant l'unique vecteur propre associé à la valeur propre 1, $|\psi_0\rangle \equiv |\varpi_1\rangle$, et (21.3) s'écrit aussi :

$$p_1(t) = |\langle \psi_0 | U(t) | \psi_0 \rangle|^2 \ . \tag{21.5}$$

Supposons maintenant que l'instant t de mesure soit très petit comparé aux inverses de toutes les fréquences de Bohr ; alors, toutes les exponentielles apparaissant dans le développement de $U(t)|\psi_0\rangle$ sur les états propres de H peuvent être développées, et l'amplitude de transition est approximativement donnée par :

$$A_{0\to 0}(t) \equiv \langle \psi_0 | U(t) | \psi_0 \rangle \simeq \langle \psi_0 | 1 + \frac{Ht}{\mathrm{i}\hbar} + \frac{1}{2!}\left(\frac{Ht}{\mathrm{i}\hbar}\right)^2 | \psi_0 \rangle + \mathcal{O}(t^3) \ . \tag{21.6}$$

Dans ces conditions :

$$p_1(t) = |A_{0\to 0}(t)|^2 \simeq \left(1 - \frac{t^2}{2\hbar^2}\langle H^2\rangle\right)^2 + \frac{t^2}{\hbar^2}\langle H\rangle^2 \simeq 1 - \hbar^{-2}\Delta H^2 t^2 + \mathcal{O}(t^4) \ . \tag{21.7}$$

Toutes les moyennes $\langle \ldots \rangle$ sont prises avec l'état $|\psi_0\rangle$, ΔH^2 est l'écart quadratique[4] de l'énergie dans l'état $|\psi_0\rangle$: $\Delta H^2 \stackrel{\text{déf}}{=} \langle H^2\rangle - \langle H\rangle^2$. À la réflexion, $p_1(t)$ apparaît clairement comme une probabilité *conditionnelle* : dire que l'état de départ est $|\psi_0\rangle$ est équivalent à affirmer que l'on a fait juste auparavant une mesure qui a produit le résultat $\varpi_\lambda = 1$; avec la notation habituelle pour une probabilité conditionnelle, on peut donc écrire :

$$p_1(t) \equiv p(t|0) \ , \tag{21.8}$$

et $p(t|0)$ est bien la probabilité de trouver 1 à l'instant t *sachant* que l'on a trouvé 1 à $t = 0$. En disant les choses autrement : supposons que l'on fasse deux mesures successives de l'observable P_0 aux deux instants t_1 et $t_2 > t_1$. Si on trouve la valeur 1 en t_1 (mesure *positive*), alors, l'état du système à $t_1 + 0$ est encore $|\psi_0\rangle$; l'état à l'instant t_2 est donc $U(t_2 - t_1)|\psi_0\rangle$. Dans ces conditions, la probabilité de trouver à nouveau la valeur propre 1 à l'instant t_2 sachant que l'on a trouvé 1 à l'instant t_1 est bien :

$$p(t_2|t_1) = p_1(t_2 - t_1) \tag{21.9}$$

soit approximativement, d'après (21.7) :

$$\boxed{p(t_2|t_1) \simeq 1 - \frac{\Delta H^2}{\hbar^2}(t_2 - t_1)^2} \tag{21.10}$$

[4]Tout ce qui suit suppose évidemment que l'écart-type ΔH de l'énergie est *fini* ; il en va toujours ainsi quand le spectre d'énergie est borné supérieurement. Dans le cas contraire, il faut que l'état initial ne soit pas trop dispersé sur l'ensemble des états propres de H et/ou que les énergies propres n'augmentent pas trop vite avec les nombres quantiques – autrement, ΔH est donné par une série divergente. La divergence des fluctuations est révélateur d'une distribution de probabilité dite *large* (loi de Cauchy (lorentzienne), loi de Pareto, etc.), donnant une grande pertinence aux événements *rares*.

Cette probabilité ne dépend que de la différence des temps, puisque H est depuis le début supposé indépendant du temps.

Effectuons maintenant une série de N mesures de P_0 à des instants successifs *équidistants* $t_n = n\Delta t$, $n = 1, 2, \ldots, N$, où $\Delta t = \frac{t}{N}$, t et N étant fixés. À l'issue de cette série, la probabilité d'avoir trouvé à chaque fois la valeur propre 1 (toutes les mesures donnent un résultat positif) est[5] :

$$p_{\mathrm{pos}}(t_1, t_2, \ldots, t_N) = p(t_N|t_{N-1})\, p(t_{N-1}|t_{N-2}) \ldots p_1(t_2|t_1) \simeq \left(1 - \frac{\Delta H^2}{\hbar^2}\frac{t^2}{N^2}\right)^N . \tag{21.11}$$

En utilisant maintenant :

$$\left(1 - \frac{C}{N^2}\right)^N = \mathrm{e}^{N \ln\left(1 - \frac{C}{N^2}\right)} \simeq \mathrm{e}^{-\frac{C}{N}} \simeq 1 \qquad (N \gg 1) , \tag{21.12}$$

on voit que la probabilité p_{pos} que *toutes* les mesures donnent un résultat positif tend vers 1 quand le nombre de mesures tend vers l'infini :

$$\boxed{\lim_{N\to\infty} p_{\mathrm{pos}}(t_1, t_2, \ldots, t_N) = 1} \tag{21.13}$$

Autrement dit, un système observé (mesuré) en “continu” n'évolue pas.

♦ *Remarques*

1. La probabilité de trouver toujours la même valeur (1 en l'occurrence) lors de la mesure de la *variable binaire* associée à P_0 tend donc vers 1 quand le nombre de mesures en un temps fini augmente indéfiniment. Ceci est complètement différent de la probabilité de trouver toujours pile lors du jet d'une pièce, effectué un grand nombre de fois, N : cette dernière vaut 2^{-N} et tend vers zéro exponentiellement. Fondamentalement, la différence tient à la corrélation statistique des événements, forte pour l'effet Zénon, nulle (ou supposée négligeable) pour le jet de la pièce. Pour l'effet Zénon quantique, plus on raccourcit l'intervalle de temps entre deux mesures successives, plus grande est la probabilité de retrouver la même valeur que lors de la mesure immédiatement précédente. Pour le jet de la pièce, tout nouveau lancement est supposé complètement indépendant de tous les précédents.

2. L'instabilité supposée de l'état $|\psi_0\rangle$ ne joue aucun rôle et n'a été adoptée que pour fixer les idées. La même conclusion peut être tirée pour un système à deux niveaux séparés par l'énergie $\hbar\omega_0$, dont la dynamique libre est périodique de période $\frac{2\pi}{\omega_0}$; en la circonstance, la seule condition à retenir est $\Delta t \ll \omega_0^{-1}$.

[5] La probabilité des N résultats positifs est le produit des probabilités de transition $p(t_{n+1}|t_n)$: il s'agit donc d'une *chaîne de Markov*, caractérisée par le fait que l'évolution à partir de l'instant qui suit immédiatement une mesure ne dépend que de l'instant présent ; c'est une conséquence du fait que l'équation de Schrödinger est une équation aux dérivées partielles du *premier* ordre en temps.

3. La conclusion exprimée par (21.13) repose crucialement sur le fait que la probabilité $p_1(t)$ s'écarte *quadratiquement* de la valeur 1 (voir (21.7)). Si la variation était linéaire (par exemple, le début d'une exponentielle décroissante), $p_1(t) \simeq (1 - \Gamma t)$, on aurait :
$$p_{\text{pos}}(t_1, t_2, \ldots, t_N) \simeq \left(1 - \Gamma\frac{t}{N}\right)^N \overset{N \gg 1}{\simeq} \mathrm{e}^{-\Gamma t} \ . \tag{21.14}$$
Plus généralement, avec un comportement du type $p_1(t) \simeq (1 - Ct^\alpha)$, on obtient :
$$p_{\text{pos}}(t_1, t_2, \ldots, t_N) \simeq \left[1 - C\left(\frac{t}{N}\right)^\alpha\right]^N \overset{N \gg 1}{\simeq} \mathrm{e}^{-Ct^\alpha N^{1-\alpha}} \ . \tag{21.15}$$
Dès que $\alpha > 1$, il y a effet Zénon : c'est bien quand le système s'éloigne *lentement* de son point de départ que celui-là peut survenir. Si on considère un atome à deux niveaux $|\mathrm{e}\rangle$ (excité, radiatif) et $|\mathrm{g}\rangle$ (fondamental), la première phase de déclin de l'état excité est essentiellement exponentielle[6] sauf sur un très petit intervalle, où le comportement – parabolique comme en (21.7) – permet l'effet Zénon ; ce régime étant de courte durée, il est très difficile d'observer l'effet Zénon dans une configuration aussi simple, d'où la proposition de Cook [97] d'utiliser un atome à trois niveaux (diagramme en "V", voir p. 929). ♦

L'argument de principe ci-dessus, mettant en évidence l'effet Zénon sur le plan des concepts, est fondé sur la définition idéale de l'opération de mesure en Mécanique quantique, caractérisée notamment par la réduction instantanée du paquet d'ondes, et sans se risquer à introduire "l'appareil de mesure" dans le formalisme (le peut-on, d'ailleurs ?). L'observation de l'effet est liée à la construction d'un scénario permettant de réaliser expérimentalement une suite de processus singeant aussi bien que possible l'opération de mesure au sens de la Mécanique quantique.

Le ralentissement d'une dynamique quantique se produit dans une grande variété de contextes, sans qu'il y ait toujours besoin de faire référence à la notion de mesure. Soit par exemple un système à *petit* nombre de niveaux couplé à un grand système, ce dernier étant distingué par le fait qu'il possède au contraire un très grand nombre de degrés de liberté. Cette caractéristique se traduit usuellement par la propriété pour le spectre d'énergie du grand système de former un quasi-continuum. Dès lors, il existe un grand nombre de couples de fréquences dont la différence est très petite, associés à des périodes très longues ; au total, les temps de récurrence (ou de pseudo-récurrences) deviennent gigantesques au point de perdre tout sens pour le physicien : l'irréversibilité de la dynamique du petit système couplé au grand devient un état de fait (voir par

[6]Toutefois, aux temps très longs, le déclin suit une loi puissance $\sim t^{-\lambda}$, qui est *relativement* dominante une fois l'exponentielle éteinte. L'analyse asymptotique de l'amplitude $A(t) = \langle\Psi(t)|\Psi(0)\rangle$ montre que $A(t) \sim t^{-3}$ (dans $\mathbb{R}^3$). Ce résultat dépend fortement de la loi de dispersion des particules émises, ici des photons, dont la loi de dispersion $E = \hbar kc$ est *linéaire* en k. Pour une particule de masse non-nulle, $E = \sqrt{(\hbar kc)^2 + m^2c^4}$, la même analyse donne $\lambda = 3/2$. Tout vient du fait que ces lois de dispersion apparaissent dans des dénominateurs d'énergie et que le comportement aux grands temps est piloté par les fréquences "infra-rouges", c'est-à-dire par les petites valeurs de k. Pour une dispersion linéaire, c'est donc k^{-1} qui est pertinent, dans le second cas, c'est la constante mc^2. Au total, le comportement en $k \sim 0$ de l'intégrand à analyser est franchement différent, d'où les deux valeurs obtenues pour l'exposant λ.

exemple le problème 14.6.11 du Tome I). Tout comme en Physique classique, on peut parler de *temps de Poincaré*, qui diverge avec le nombre de degrés de liberté[7].

Lorsque le couplage entre les deux systèmes est faible en un sens à préciser, il existe des approximations légitimes montrant que sur un grand intervalle de temps, la dynamique oscillante est amortie suivant un simple régime exponentiel (voir par exemple l'approximation dite *du pôle*, p. 1167). Cette conclusion repose sur le fait que le couplage avec le grand système décale les pôles de la résolvante[8] de l'axe réel vers le demi-plan complexe *inférieur* (plus précisément dans le deuxième feuillet de Riemann), fabriquant ce que l'on appelle dans d'autres contextes une *résonance*, d'autant plus fine que le temps de relaxation du petit système est long – dans la limite où ce temps est infiniment long, on obtient une résonance de largeur nulle, caractéristique d'une durée de vie infiniment grande : si $\hbar\Gamma$ est largeur de la résonance, le temps de relaxation (ou durée de vie de l'état instable) est $\sim \Gamma^{-1}$. Outre l'apparition de la largeur, liée à la partie imaginaire de ce qui joue le rôle d'une valeur propre de H à un niveau élémentaire, la partie réelle du pôle est (un peu) différente de l'énergie "nue" : on parle de *renormalisation* de l'énergie, et c'est ainsi que l'on peut rendre compte du *Lamb shift* d'un atome (voir p. 898) ; dans ce dernier cas, le grand système est le champ électromagnétique quantifié – c'est pourquoi on parle alors de *déplacement radiatif*. Le *Lamb shift* résulte des fluctuations du vide électromagnétique puisque, comme tous les objets quantiques, les champs ont des *fluctuations* : dans l'état vide (zéro photon *en moyenne*), les valeurs moyennes des champs sont certes nulles, mais les fluctuations quantiques irréductibles demeurent inéluctablement.

D'un autre point de vue, pragmatique, le comportement exponentiel d'une relaxation peut être d'emblée fabriqué de façon purement phénoménologique en incorporant des parties imaginaires (négatives) dans les énergies, ce qui revient peu ou prou à introduire un Hamiltonien *ad hoc* non hermitique. Cette recette, validée par l'approximation du pôle, permet de faire l'économie de la description quantique du grand système, forcément un peu laborieuse ; le problème 21.6.1 traite en détail le cas d'un atome à trois niveaux Cook [97] a proposé une expérience destinée à mettre en évidence l'effet Zénon, par une inhibition de l'absorption induite dans un atome à trois niveaux, du type de celui considéré dans le problème sus-visé. Avant d'aborder la discussion de l'expérience décisive d'Itano et al. [98], et pour en comprendre le traitement théorique qu'en ont proposé ces auteurs, quelques éléments concernant l'opérateur densité sont nécessaires.

---

[7]La divergence du temps de Poincaré pour un système macroscopique est l'une des clés de voûte de la Mécanique statistique (classique) pour la justification théorique du Second principe de la Thermodynamique. De façon *duale*, le temps de *décohérence* d'un système macroscopique est infiniment court, un résultat majeur de la théorie de la décohérence (voir section 21.4), laquelle vise à expliquer pourquoi le monde perceptible avec les sens communs est d'apparence classique.

[8]Essentiellement (voir chapitre 23, sous-section 23.2.3), la résolvante est la transformée de Laplace (ou de Fourier) de l'opérateur d'évolution $U(t)$. C'est un avatar de fonction de Green, dont les singularités dans $\mathbb{C}$ donnent le spectre de H et déterminent donc la dynamique (pour plus de détails, voir [15] et plus loin, chapitre 24). En particulier, quand le spectre de H contient une composante continue, la résolvante possède des points de branchement ; la (ou les) coupure correspondante définit la frontière commune des différents feuillets constituant la surface de Riemann – voir l'exemple du traitement de l'émission spontanée, chapitre 25, section 25.2.

L'opérateur densité

Classiquement, et qu'il s'agisse de l'état d'équilibre ou non, le passage de la Mécanique à la Mécanique statistique s'opère par l'introduction de la fonction de densité $\rho_{\text{cl}}(\{q\}, \{p\}, t)$ qui donne la densité de probabilité dans l'espace des phases du système, une fois défini l'ensemble de Gibbs approprié (voir par exemple [99]). À cette fonction classique correspond en Mécanique statistique quantique un *opérateur densité*, qui en est la quantité centrale tout comme l'est la fonction d'onde pour la Mécanique quantique. Mécanique et Statistique quantiques se fondent l'une dans l'autre à température nulle, tout comme leurs homologues classiques.

Pour un système quantique d'états $|E_n\rangle$ à l'équilibre thermique avec un thermostat à température T (système *canonique*), les probabilités de Boltzmann $P_n(\beta)$ sont proportionnelles à $\mathrm{e}^{-\beta E_n}$, la décroissance exponentielle avec l'énergie venant du fait que, comparé au système étudié, le thermostat (*réservoir*) est aussi grand que l'on veut (c'est lui qui fixe la température) ; adoptant la numérotation conventionnelle $E_n \leq E_{n+1}$, $P_n \geq P_{n+1}$. Dans cette situation, on définit l'opérateur densité[9] (d'équilibre) $\rho(\beta)$:

$$\rho(\beta) \stackrel{\text{déf}}{=} C\,\mathrm{e}^{-\beta H} \ , \qquad \beta \stackrel{\text{déf}}{=} \frac{1}{k_{\mathrm{B}}T} \tag{21.16}$$

où C est une constante de normalisation ; $\rho(\beta)$ est visiblement hermitique. Sur la base des états propres du Hamiltonien du système, ρ est représenté par la matrice diagonale dont les éléments de matrice sont les probabilités de Boltzmann :

$$\rho(\beta) = \sum_n C\mathrm{e}^{-\beta E_n}|E_n\rangle\langle E_n| \equiv \sum_n \rho_{nn}|E_n\rangle\langle E_n| \ ; \tag{21.17}$$

la normalisation des probabilités s'écrit $C\sum_n \mathrm{e}^{-\beta E_n} = 1$, soit $C = Z^{-1}(\beta)$ où :

$$Z(\beta) \stackrel{\text{déf}}{=} \sum_n \mathrm{e}^{-\beta E_n} \ ; \tag{21.18}$$

avec cette définition, l'opérateur densité canonique normalisé s'écrit :

$$\boxed{\rho(\beta) = \frac{1}{Z(\beta)}\sum_n \mathrm{e}^{-\beta E_n}\,|E_n\rangle\langle E_n| \qquad \mathrm{Tr}\,\rho(\beta) = 1} \tag{21.19}$$

$Z(\beta)$ est la *fonction de partition* (canonique) ; cette fonction est au cœur de toutes les relations thermodynamiques et permet de trouver toutes les valeurs moyennes dans la situation canonique. Ainsi, la moyenne de l'énergie est $\langle E\rangle = Z^{-1}\sum_n E_n\mathrm{e}^{-\beta E_n}$, soit :

$$\langle E\rangle = -\frac{\partial}{\partial\beta}\ln Z(\beta) \tag{21.20}$$

[9]Remarquer que ρ a la même forme que le propagateur $U(t)$ pour un système dont le Hamiltonien H est indépendant du temps – très précisément $\rho(\beta) = U(t = -\mathrm{i}\hbar\beta)$ (on dit que l'on passe en *temps imaginaire*). De même, on voit que $\partial_\beta\rho = -H\rho$, équation dite *de Bloch* qui, *mutatis mutandis*, coïncide avec l'équation de Schrödinger pour $U(t)$ (voir problème 23.4.15). Cette identité formelle est l'une des raisons pour lesquelles on retrouve très souvent les mêmes techniques de calcul en Mécanique statistique ($T > 0$) et en Mécanique quantique ($T = 0$) (méthodes perturbatives et variationnelles, intégrale de Feynman, etc.).

et de même pour toute autre moyenne, premièrement exprimée en termes des probabilités de Boltzmann. Bien évidemment, la représentation matricielle de l'opérateur densité dépend de la base utilisée ; si $\{a_n\}_n$ est la base propre (orthonormalisée) d'une certaine observable A, $A|a_m\rangle = a_m|a_m\rangle$, on peut écrire :

$$\rho(\beta) = Z^{-1}(\beta) \sum_{mm'} |a_m\rangle\langle a_m|\mathrm{e}^{-\beta H}|a_{m'}\rangle\langle a_{m'}| \ . \tag{21.21}$$

Sur cette base, la matrice de ρ n'est plus diagonale, sauf bien sûr si $[A,\, H] = 0$.

L'entropie statistique canonique est $S(\beta) = -k_\mathrm{B} \sum_n P_n(\beta) \ln P_n(\beta)$ (voir [99], par exemple), une fonction positive puisque $0 \leq P_n \leq 1$ quel que soit n ; si on utilise la base propre du Hamiltonien, elle peut s'écrire comme :

$$S(\beta) = -k_\mathrm{B} \sum_n \rho_{nn}(\beta) \ln \rho_{nn}(\beta) \ , \qquad \rho_{nn}(\beta) = Z^{-1}(\beta)\, \mathrm{e}^{-\beta E_n} \ ; \tag{21.22}$$

cette expression peut être lue comme une certaine trace, donnant finalement la forme compacte :

$$\boxed{S(\beta) = -k_\mathrm{B} \mathrm{Tr}\, [\rho(\beta) \ln \rho(\beta)]} \tag{21.23}$$

En tant que trace, c'est un invariant dans tout changement de base. Noter que l'entropie s'annule si $\rho_{nn} \rightarrow \delta_{nn_0}$, une situation obtenue dans la limite de température nulle :

$$\lim_{\beta \rightarrow +\infty} \rho(\beta) = |E_1\rangle\langle E_1| \ , \tag{21.24}$$

où $|E_1\rangle$ est l'état fondamental, $E_1 = \inf_n E_n$; c'est la situation d'équilibre de la Mécanique quantique (système dans l'état fondamental) : à température nulle, les Mécaniques quantique et statistique (quantique) contiennent bien strictement la même information. Dans cette limite, ρ est *idempotent* (égal à ses puissances entières positives) :

$$\boxed{\rho^2(+\infty) = \rho(+\infty) \qquad \text{(cas pur)}} \tag{21.25}$$

et on dit alors qu'il s'agit d'un *cas pur* ($S = 0$). Au contraire, à toute température finie :

$$\boxed{\rho^2(\beta) \neq \rho(\beta) \qquad \text{(cas mixte)}} \tag{21.26}$$

et on parle alors de *cas mixte* ($S > 0$) ; après avoir défini une relation d'ordre évidente entre *opérateurs*, on peut écrire $\rho^2(\beta) < \rho(\beta)$ puisque $\rho_{nn}^2 < \rho_{nn}\ \forall T > 0$ si $\rho_{nn} \neq \delta_{nn_0}$.

Toutes les définitions précédentes peuvent en réalité être écrites formellement sans faire référence à la situation canonique d'équilibre, introduite ci-dessus pour mettre en évidence l'origine physique de l'opérateur ρ dont il est visible qu'il constitue une généralisation de la description de l'état quantique d'un système, relativement à celle que pourvoit un vecteur d'état $|\Psi\rangle$. Une base de représentation étant choisie, celle d'une observable A, on peut en toute généralité *définir* un certain opérateur ρ diagonal :

$$\boxed{\rho \stackrel{\text{déf}}{=} \sum_n \rho_{nn}\, |a_n\rangle\langle a_n|} \tag{21.27}$$

dont tous les éléments de matrice diagonaux sont positifs ou nuls, de somme calée à 1 (ceux-ci étant réels, ρ est hermitique) ; ces propriétés assurent que chaque ρ_{nn} peut être interprété comme une certaine probabilité P_n, $P_n \stackrel{\text{déf}}{=} \rho_{nn}$. Par définition, un système décrit par l'opérateur densité (21.27) est tel que la mesure de l'observable A donne la valeur a_n avec la probabilité P_n :

$$\boxed{\text{Système décrit par } \rho = \sum_n \rho_{nn}\,|a_n\rangle\langle a_n| \iff \text{Prob[mesure de } A \to a_n] = \rho_{nn}} \tag{21.28}$$

Avec les P_n, on peut construire une fonction positive du genre entropie[10], $\mathcal{S}$, définie à un facteur conventionnel près, $\mathcal{S}_0 > 0$:

$$\mathcal{S} \stackrel{\text{déf}}{=} -\mathcal{S}_0 \sum_n P_n \ln P_n \ . \tag{21.29}$$

L'expression (21.27) montre[11] que ρ est idempotent ssi $\rho_{nn} = \delta_{nn_0}$, donnant $\mathcal{S} = 0$, et c'est ainsi que l'on définit en toute généralité un cas pur :

$$\text{cas pur} \iff \rho_{nn} = \delta_{nn_0} \iff \rho^2 = \rho \tag{21.30}$$

En pareil cas, (21.27) montre que ρ est égal à $|a_{n_0}\rangle\langle a_{n_0}|$; l'observable A étant quelconque, on voit finalement qu'un cas pur correspond au cas où il existe un certain *ket* normalisé $|\psi\rangle$ tel que l'opérateur densité peut s'écrire $\rho = |\psi\rangle\langle\psi|$:

$$\boxed{\text{cas pur} \iff \exists\ |\psi\rangle,\ \rho = |\psi\rangle\langle\psi| \iff \rho^2 = \rho \iff \mathcal{S} = 0} \tag{21.31}$$

L'idempotence de ρ est évidente sur cette écriture, puisque $|\psi\rangle$ est supposé normalisé : $\rho^2 = |\psi\rangle\langle\psi||\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \rho$. Le cas $|\psi\rangle = |E_n\rangle$ correspond à un certain état stationnaire ; s'il s'agit plus précisément de l'état fondamental, c'est aussi la limite température nulle de l'opérateur canonique.

En raison des définitions précédentes, la valeur moyenne[12] d'une observable A

[10] Si on choisit $\mathcal{S}_0 = (\log_2 \mathrm{e})^{-1}$, en relation avec le *bit binaire* contituant l'*atome* d'information, la fonction $\mathcal{S}$ coïncide avec la fonction dite *information manquante* $I \stackrel{\text{déf}}{=} -\sum_n P_n \log_2 P_n$ définie par Shannon et Wiener en Théorie de l'information (elle aussi appelée *entropie*) ; si n est borné (nombre fini d'*états*, ou d'*événements élémentaires*), $1 \leq n \leq N$, cette fonction a toutes les propriétés souhaitables (et en fait requises en préalable pour assurer l'unicité). Notamment, elle est bornée inférieurement par 0, valeur atteinte si toutes les probabilités sont nulles sauf une ($P_n = \delta_{nn_0}$, – situation "pure", non "désordonnée"), et supérieurement par $\log_2 N$, quand $P_n = 1/N$ quel que soit n (équiprobabilité, situation la plus "désordonnée"). Pour une introduction très claire, voir l'ouvrage de Rényi [100].

On notera la proximité conceptuelle avec le postulat fondamental de la Mécanique statistique [99] (*"Tous les états de même énergie ont la même probabilité"*), assurant *ipso facto* que l'entropie (statistique) est *maximale*.

[11] Pour un système à N états, il suffit d'écrire $\rho_{nn} = (1-\varepsilon)\delta_{nn_0} + \varepsilon(1-\delta_{nn_0})/(N-1)$ et de prendre la limite $\varepsilon \to 0$.

[12] Rappelons que l'expression *valeur moyenne dans l'état...* est une formule concise pour dire que si l'on effectue un grand nombre N_{exp} de mesures de A sur le système préparé à chaque fois dans l'état en question, alors la moyenne statistique des résultats enregistrés tend vers $\langle A\rangle$ quand $N_{\text{exp}} \to \infty$. En dehors de toute opération de mesure, on ne sait pas ce que "vaut" l'observable A, et encore moins ce qu'est sa "moyenne" (?) sur des opérations non encore effectuées.

dans l'état décrit par l'opérateur (21.27) est l'espérance mathématique :

$$\langle A \rangle = \sum_n P_n a_n \equiv \sum_n \rho_{nn} a_n \ , \tag{21.32}$$

une expression qui, à nouveau, peut s'écrire de façon compacte en introduisant la trace :

$$\boxed{\langle A \rangle = \text{Tr }(A\rho) = \text{Tr }(\rho A)} \tag{21.33}$$

et montre l'invariance de la valeur moyenne par rapport à une représentation choisie.

La structure algébrique de la matrice de ρ sur une base donnée est *qualitativement* différente selon que l'on est dans un cas pur ou dans un cas mixte – ce que l'on peut soupçonner au seul fait que ρ est idempotent dans le premier cas, non idempotent dans le second. Pour un cas pur, $\exists\, |\psi\rangle$, $\rho = |\psi\rangle\langle\psi|$; décomposant alors $|\psi\rangle$ sur la base propre de A, $|\psi\rangle = \sum_n c_n |a_n\rangle$, il vient :

$$\rho = \Big(\sum_n c_n |a_n\rangle\Big)\Big(\sum_m c_m^* \langle a_m|\Big) = \sum_{nm} c_n c_m^* |a_n\rangle\langle a_m| \ , \tag{21.34}$$

écriture où l'on voit que, pour un cas pur, ρ_{nm} est le *produit* de deux nombres $c_n c_m^*$ contraints par $\sum_n |c_n|^2 = 1$:

$$\boxed{\text{cas pur} \quad \Longleftrightarrow \quad \rho_{nm} = c_n c_m^*} \tag{21.35}$$

Cette propriété de *factorisation* de tous les éléments de matrice est indépendante de la représentation choisie ; sur la base propre d'une autre observable B ne commutant pas avec A ($B|b_p\rangle = b_p|b_p\rangle$, $|\psi\rangle = \sum_p d_p|b_p\rangle$), ρ est représenté par une autre matrice, dont les éléments sont égaux à $d_p d_q^*$: s'ils sont effectivement différents des $c_n c_m^*$, ils se présentent encore sous la forme factorisée.

Au contraire, pour un cas mixte, ρ est de la forme (21.27) ; en introduisant les états propres de B, on a :

$$\rho = \sum_{pq} \Big(\sum_n \rho_{nn} \langle b_p|a_n\rangle\langle a_n|b_q\rangle\Big) |b_p\rangle\langle b_q| \ ; \tag{21.36}$$

la quantité entre parenthèses est l'élément de matrice $\langle b_p|\rho|b_q\rangle$, qui n'est pas un produit de deux facteurs – sauf si, à nouveau, $\rho_{nn} = \delta_{nn_0}$, auquel cas ρ est le cas pur $|a_{n_0}\rangle\langle a_{n_0}|$, dont les éléments non-diagonaux sur la base propre de B sont égaux à $\langle b_p|a_{n_0}\rangle\langle a_{n_0}|b_q\rangle$, qui est bien un produit de deux facteurs.

Les définitions précédentes, énoncées d'abord à propos de l'équilibre thermodynamique, puis généralisées formellement, peuvent manifestement être appliquées à une situation variable dans le temps, puisqu'aucune référence au temps n'a été faite. Décrire la dynamique passe de prime abord par l'écriture de l'équation d'évolution pour l'opérateur $\rho(t)$, qui généralise au champ statistique l'équation de Schrödinger pour $|\Psi(t)\rangle$ et permet (en principe) de trouver, par exemple, la *relaxation* d'un système préparé dans un état hors d'équilibre.

Dans un cadre classique et pour une situation hors d'équilibre, la manipulation de l'équation fondamentale de la dynamique – ou ses variantes plus élaborées, les équations canoniques de Hamilton – fournit l'équation :

$$\boxed{\frac{\partial \rho_{\rm cl}}{\partial t} = \{H,\, \rho_{\rm cl}\}} \tag{21.37}$$

où $\{H,\, \rho_{\rm cl}\}$ est le crochet de Poisson : c'est l'équation de Liouville (classique), traduction de l'équation de conservation de la probabilité $\partial_t \rho + {\rm div}\vec{j} = 0$, $\vec{j}$ étant le vecteur courant dans l'espace des phases (le fluide de probabilité s'écoule sans pertes). Comme toujours, la version quantique de l'équation classique s'obtient en remplaçant le crochet de Poisson par le bon commutateur impliquant les observables associées, $({\rm i}\hbar)^{-1}[\ ,\ ]$, donnant ainsi l'équation de Liouville quantique[13] :

$$\boxed{{\rm i}\hbar\frac{\partial \rho}{\partial t} = [H,\, \rho]} \tag{21.38}$$

C'est cette équation que l'on postule en Mécanique statistique quantique – tout comme on postule l'équation de Schrödinger ${\rm i}\hbar\frac{\partial \Psi}{\partial t} = H\Psi$; quand H ne dépend pas du temps, son intégration formelle donne :

$$\boxed{\rho(t) = {\rm e}^{\frac{1}{{\rm i}\hbar}Ht}\,\rho(0)\,{\rm e}^{-\frac{1}{{\rm i}\hbar}Ht}} \tag{21.39}$$

Quand l'état initial est diagonal sur la base propre du Hamiltonien, qu'il s'agisse d'un cas pur ou d'un cas mixte, alors $[\rho(0),\, H] = 0$, de sorte que $\rho(t) = \rho(0)$: le système n'évolue pas au cours du temps (état d'équilibre).

Si l'état de départ est un cas pur, il existe un vecteur $|\psi\rangle$ tel que $\rho(0) = |\psi\rangle\langle\psi|$, vecteur qui est visiblement l'état initial $|\Psi(t=0)\rangle$; alors :

$$\rho(t) = {\rm e}^{\frac{1}{{\rm i}\hbar}Ht}\,|\psi\rangle\langle\psi|\,{\rm e}^{-\frac{1}{{\rm i}\hbar}Ht} \equiv |\Psi(t)\rangle\langle\Psi(t)| \ . \tag{21.40}$$

D'où le résultat important : *un cas pur reste un cas pur si l'évolution se fait selon l'équation de Schrödinger*[14]. Aucune évolution hamiltonienne (unitaire) ne peut transformer un cas pur en cas mixte ou inversement – tant qu'il est question de l'opérateur densité total[15], c'est-à-dire décrivant le système (isolé) dans sa globalité. Par ailleurs, (21.39) montre que $\rho(t)$ est un opérateur hermitique si $\rho(t=0)$ l'est, et que la trace de $\rho(t)$ est une constante du mouvement (la trace est invariante par permutation circulaire) :

$$\boxed{{\rm Tr}\ \rho(t) = {\rm Tr}\ \rho(0) = 1} \tag{21.41}$$

[13] *Au signe près*, cette équation est identique à l'équation de Heisenberg ${\rm i}\hbar\partial_t A_{\rm H}(t) = [A_{\rm H}(t),\, H]$ pour l'image de Heisenberg de l'observable A indépendante du temps dans l'image de Schrödinger (voir Tome I, sous-section 14.1.2).

[14] On pressent que, justement, le comportement *hamiltonien* exclut radicalement tout processus de mesure au sens quantique ; ceci se confirmera peu après, voir (21.48).

[15] Au contraire, des traces partielles sur certains degrés de liberté peuvent, moyennant (éventuellement) les bonnes limites du genre "limite thermodynamique", produire une équation apte à décrire le basculement cas pur → cas mixte au cours du temps. On connaît des scénarios analogues dans le cadre classique (voir note 18, p. 927) pour produire l'irréversibilité temporelle dans un formalisme pourtant strictement mécanique.

équation ne traduisant rien d'autre que la conservation de la somme des probabilités.

Le sens physique de $\rho(t)$ est le même que celui de l'opérateur d'équilibre. Notamment, la valeur moyenne de toute observable s'obtient encore par la formule (21.33), pour l'établissement de laquelle aucune référence au temps n'a été faite :

$$\boxed{\langle A\rangle(t) = \mathrm{Tr}\,[A\rho(t)] = \mathrm{Tr}\,[\rho(t)A]} \tag{21.42}$$

La notation pour $\langle A\rangle(t)$ rappelle le fait que la dépendance temporelle vient de la prise de moyenne.

Sur la base propre du Hamiltonien, $\{|E_n\rangle\}_n$, $\rho(t)$ est représenté par une matrice en général non identiquement diagonale (quand elle est diagonale à tout instant, ρ est en fait indépendant du temps). Sur cette base, l'équation de Liouville prend la forme d'un système linéaire :

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle E_n|\rho(t)|E_m\rangle = (E_m - E_n)\langle E_n|\rho(t)|E_m\rangle \iff \rho_{nm}(t) = \mathrm{e}^{\mathrm{i}(\omega_n-\omega_m)t}\,\rho_{nm}(0) \tag{21.43}$$

($\omega_n \stackrel{\text{déf}}{=} \hbar^{-1}E_n$). Les éléments non-diagonaux oscillent donc aux fréquences de Bohr quant aux éléments diagonaux, ils sont indépendants du temps et donnent les probabilités pour qu'une mesure effectuée à l'instant t donne la valeur E_n pour l'énergie :

$$\rho_{nn}(t) \equiv \langle E_n|\rho(t)|E_n\rangle = \rho_{nn}(0) = \mathrm{Prob}[\text{mesure de } E \rightarrow E_n,\, t] \ ; \tag{21.44}$$

la constance en temps des éléments diagonaux sur la base propre de H n'est qu'un avatar de la conservation de l'énergie. En définitive, les éléments diagonaux de ρ sont ici les *populations* (relatives) des différents états propres du Hamiltonien au sein de l'ensemble statistique.

En tant qu'opérateur représenté par une matrice sur une base de représentation, celle du Hamiltonien par exemple, ρ possède des éléments non-diagonaux que Cohen-Tannoudji a proposé d'appeler *cohérences*, en raison de leur rôle essentiel dans les *interférences* quantiques. En effet, soit un cas pur construit avec l'état non-stationnaire[16] :

$$|\Psi(t)\rangle = c_1\mathrm{e}^{-\mathrm{i}\omega_1 t}\,|E_1\rangle + c_2\mathrm{e}^{-\mathrm{i}\omega_2 t}\,|E_2\rangle \ ; \tag{21.45}$$

l'opérateur densité correspondant est $\rho(t) = |\Psi(t)\rangle\langle\Psi(t)|$ soit, après distribution :

$$\rho(t) = |c_1|^2|E_1\rangle\langle E_1| + |c_2|^2|E_2\rangle\langle E_2| + c_1c_2^*\mathrm{e}^{-\mathrm{i}(\omega_1-\omega_2)t}|E_1\rangle\langle E_2| + c_2c_1^*\mathrm{e}^{-\mathrm{i}(\omega_2-\omega_1)t}|E_2\rangle\langle E_1| \ , \tag{21.46}$$

ou, sous forme matricielle :

$$\rho(t) = \begin{bmatrix} |c_1|^2 & c_1c_2^*\mathrm{e}^{-\mathrm{i}(\omega_1-\omega_2)t} \\ c_2c_1^*\mathrm{e}^{-\mathrm{i}(\omega_2-\omega_1)t} & |c_2|^2 \end{bmatrix} \equiv \begin{bmatrix} \rho_{11} & \rho_{12}(t) \\ \rho_{21}(t) & \rho_{22} \end{bmatrix} \ . \tag{21.47}$$

Les cohérences $\rho_{nm}(t)$, $n \neq m$, varient harmoniquement à la pulsation de Bohr, laquelle caractérise l'oscillation de la valeur moyenne d'une observable ne commutant pas avec le

[16] Un système à deux niveaux est choisi pour la simplicité. Bien évidemment, tout se généralise à un nombre quelconque d'états.

Hamiltonien. Elles représentent la *cohérence de phase* entre les différentes composantes d'un état non-stationnaire. En outre, leur existence – et leur forme en *produit* de deux facteurs – est essentiellement liée à l'idempotence de ρ, c'est-à-dire à la propriété de ρ de représenter un cas pur. Lorsque, dans un contexte à préciser, les cohérences ont la possibilité de *décroître* au cours du temps[17], le cas pur évolue en cas mixte.

L'équation de Liouville (21.38) engendre une évolution Hamiltonienne au sens où un cas pur le reste au cours du temps. La transition cas pur $\rightarrow$ cas mixte peut s'opérer essentiellement de deux façons : soit en appliquant, formellement et strictement, le scénario de la mesure idéale telle que le définit la Mécanique quantique – en nécessité logique et compte tenu du postulat de réduction du paquet d'ondes –, soit en couplant explicitement le sytème quantique à un grand système (ayant un nombre infini de degrés de liberté à la limite thermodynamique), appelé *réservoir*, et en intégrant sur les degrés de liberté de ce dernier (ce qui conduit à un opérateur densité *réduit*). La description de la dynamique de cet opérateur réduit peut apporter la clé d'une justification physique de l'apparence classique du monde macroscopique puisque, selon la théorie de la décohérence (voir section 21.4), le temps de décohérence d'un grand système est un infiniment petit au sens de la Physique ; les progrès en ce sens sont une avancée importante dans la compréhension physique d'un scénario théorique au demeurant fort mystérieux.

Quand on couple un système à un *réservoir*, le "petit" système reste l'objet principal d'étude et est supposé n'avoir aucune influence en retour sur la dynamique du grand système. Il en résulte que seuls les degrés de liberté du système présentent un intérêt, ceux du réservoir étant connus (d'une façon ou d'une autre et à l'échelle macroscopique), une fois pour toutes. L'idée est donc d'écrire l'équation de Liouville pour le "supersystème", puis de prendre la trace sur tous les degrés de liberté du réservoir. On obtient alors une équation pour l'opérateur densité réduit, laquelle (sauf exception) ne conserve pas le caractère pur au cours du temps[18]. Toute la difficulté est la modélisation astucieuse du réservoir et de son interaction avec le petit système, permettant au total de définir un modèle raisonnable physiquement et tractable techniquement (pour quelques exemples, voir [102] ; le modèle *spin - boson* est aussi un favori [103]). Pour ce type de problèmes, le formalisme de la fonctionnelle d'influence [104] est d'une grande utilité, autant sur le plan conceptuel que sur le plan technique.

L'opérateur densité n'est pas seulement requis pour décrire une évolution statis-

[17]Phénomène appelé *décohérence*, et présenté plus loin (section 21.4).

[18]Un phénomène analogue se produit déjà en Physique classique dans des cas très simples, même lorsque peu de degrés de liberté sont impliqués. La fonction de densité $\rho(q, p, t)$ d'un système à un degré de liberté obéit à une équation réversible par renversement du temps, puisqu'elle n'est qu'une récriture des équations de la Mécanique incorporant la conservation de la probabilité. Toutefois, la fonction réduite (marginale) $\rho_q(q, t)$, obtenue par intégration sur l'impulsion et donnant la densité de probabilité de présence de la particule au point d'abscisse q quelle que soit son impulsion p, a une évolution irréversible en général (voir par exemple l'article de Born [101], repris dans le problème 9.8.2 du Tome I). On pourra retenir l'idée suivante : prendre une trace partielle c'est perdre de l'information, une perte qui *peut* se traduire par une brisure de la symétrie de renversement du temps.

C'est aussi pour cette raison que, tout en restant strictement dans un cadre classique – et en excluant évidemment toute sensibilité aux conditions initiales –, il existe des systèmes très simples (une particule libre !) pour lesquels toute imprécision (inévitable !) sur les conditions initiales ruine toute prévision à long terme (voir Tome I, problème 9.8.2).

tique à partir d'une situation hors d'équilibre, il est aussi l'outil idéal pour la description de l'état juste avant et juste après une mesure, effectuée simultanément (et indépendamment) sur les (micro)systèmes d'un ensemble au sens statistique ; ρ est donc très bien adapté pour transcrire algébriquement l'inévitable processus de réduction du paquet d'ondes. En effet, la mesure à un instant t de l'énergie sur un ensemble décrit par ρ donne les différentes valeurs propres E_n avec des probabilités P_n, lesquelles sont précisément les éléments diagonaux ρ_{nn} de l'opérateur ρ, représenté sur la base propre de H. Juste après la mesure, une fraction $|c_1|^2 \equiv P_1$ de systèmes se retrouve dans l'état $|E_1\rangle$, une fraction $|c_2|^2 \equiv P_2$ se retrouve dans $|E_2\rangle$, etc. : l'ensemble statistique est alors décrit par un opérateur densité que la mesure a rendu *diagonal.* Revenant pour la simplicité au système à deux niveaux, la réduction du paquet d'ondes se traduit en d'autres termes par *l'annulation des cohérences* :

$$\rho_{\text{juste avant la mesure à l'instant } t} = \begin{bmatrix} \rho_{11} & \rho_{12}(t) \\ \rho_{21}(t) & \rho_{22} \end{bmatrix} \longrightarrow \begin{bmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{bmatrix} = \rho_{\text{juste après la mesure à l'instant } t} \ . \quad (21.48)$$

Cette transcription de l'opération de mesure en termes de l'opérateur densité est due à von Neumann [105]. Il vaut la peine de souligner que, si le choix d'une représentation est en général une pure affaire de commodité de calcul (seules les quantités invariantes ont un sens physique), la décision de mesurer *une observable particulière* privilégie sans discussion sa base propre, donnant à la *matrice* de $\rho(t)$ sur cette base un sens physique direct incontestable.

Clairement, la mesure transforme un cas pur en cas mixte – sauf dans le cas exceptionnel où l'état objet de la mesure est, avant cette mesure, un état propre de l'observable mesurée, ici l'énergie. On sait que, par nature, la réduction du paquet d'ondes se traduit par l'annulation des composantes du vecteur d'état orthogonales à l'état propre associé à la valeur effectivement trouvée ; transposé à l'opérateur densité, cette *projection* du vecteur d'état se concrétise par l'annulation des cohérences et la "projection" de ρ au sens où les éléments diagonaux restent ce qu'ils sont, cependant que les éléments non-diagonaux sont mis à zéro. Noter qu'il s'agit d'une opération formelle, à propos de laquelle on souhaite pouvoir imaginer un *mécanisme* physique dont cette opération est la traduction abstraite. Cette quête du Graal est l'un des enjeux de la théorie de la décohérence (voir section 21.4). ■

Revenant maintenant à l'effet Zénon, la proposition de Cook [97] d'une expérience destinée à le mettre en évidence repose sur l'idée d'une *inhibition* de l'absorption induite dans un atome à trois niveaux, notés $|\mathrm{g}\rangle$, $|\mathrm{m}\rangle$ et $|\mathrm{e}\rangle$, l'inhibition étant le symptôme du freinage de la dynamique par des mesures répétées. Par opposition avec le modèle phénoménologique analysé dans le problème 21.6.1, c'est le couple $(|\mathrm{g}\rangle, |\mathrm{m}\rangle)$ qui est sollicité par une excitation résonnante de pulsation ω et participe à l'oscillation de Rabi (voir fig. 21.1) ; la transition $|\mathrm{g}\rangle \leftrightarrow |\mathrm{m}\rangle$, d'énergie $\hbar\omega_0$, est dans le domaine des radiofréquences (quelques centaines de MHz), l'oscillation de Rabi se produit à une fréquence de l'ordre de quelques dizaines de Hertz ou même moins. Si on peut effectuer une mesure de l'état du système durant cette oscillation, on trouve soit $|\mathrm{g}\rangle$, soit $|\mathrm{m}\rangle$ – l'état $|\mathrm{e}\rangle$ étant

pour l'instant hors-circuit – avec des probabilités connues, dépendant de la position de l'instant de mesure par rapport au cycle de Rabi (et aussi de l'écart[19] à la résonance $\omega - \omega_0$). En définitive, l'opération de mesure – à définir plus précisément – révèle soit l'état $|\mathrm{g}\rangle$, soit l'état $|\mathrm{m}\rangle$.

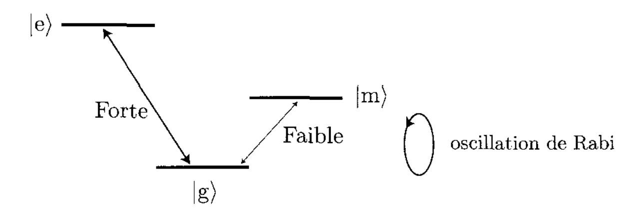


Figure 21.1: Atome à trois niveaux pour l'effet Zénon.

L'état $|\mathrm{e}\rangle$ est supposé être couplé exclusivement au fondamental $|\mathrm{g}\rangle$ par l'intermédiaire du champ électromagnétique : un *pulse* optique peut donc provoquer des transitions entre $|\mathrm{g}\rangle$ et $|\mathrm{e}\rangle$. $|\mathrm{e}\rangle$ est un état excité ordinaire, avec une durée de vie courte, 10^{-10} s pour fixer les idées.

L'atome est donc soumis d'une part à l'excitation résonnante provoquant l'oscillation de Rabi, d'autre part à un pulse optique de durée très courte, ou plusieurs *pulses* successifs, et chacun de ceux-ci provoque le *collapse* de la fonction d'onde. En effet, si l'atome est dans l'état $|\mathrm{g}\rangle$ quand arrive un tel *pulse*, des cycles d'absorption-émission spontanée de photons se produisent, se manifestant par une *diffusion* de photons dans une direction aléatoire par rapport à la direction incidente[20] ; si au contraire l'atome est dans l'état $|\mathrm{m}\rangle$, aucune diffusion de photon ne se produit. Ainsi, l'absence de photons diffusés révèle que l'état de l'atome est l'état $|\mathrm{m}\rangle$ et produit donc le *collapse* de la fonction d'onde en $|\mathrm{m}\rangle$; au contraire, la présence de photons diffusés produit le *collapse* en $|\mathrm{g}\rangle$. Pour parler de façon imagée, l'observation ou non de photons diffusés joue le rôle de *l'aiguille de l'ampèremètre.*

En outre, si on "trouve" $|\mathrm{m}\rangle$ au sens ci-dessus, l'atome reste en $|\mathrm{m}\rangle$ durant tout le pulse et un deuxième pulse rapproché en temps révèle à nouveau le même état. Si au contraire on "trouve" $|\mathrm{g}\rangle$, l'atome retourne très vite dans le même état $|\mathrm{g}\rangle$, puisque la durée de vie Γ_{e}^{-1} est très petite.

En définitive, le *pulse* optique semble bien constituer une mesure de l'état au sens où l'entend la Mécanique quantique, à un petit détail près : la durée finie du *pulse* et le séjour éphémère de l'atome dans l'état $|\mathrm{e}\rangle$ – alors que le processus de mesure, tel qu'il est ordinairement décrit, affirme l'instantanéité de la réduction du paquet d'ondes. Comme toujours, l'essentiel est que l'instantanéité supposée repose sur la considération de deux

[19] C'est seulement quand cet écart est nul que les probabilités P_{g} et P_{m} varient harmoniquement (à la pulsation Ω) entre 0 et 1 (voir problème 18.7.11 p. 773).

[20] Le terme *diffusion* est à prendre au sens de *éparpillement*, en raison de la direction aléatoire d'émission : partant d'un faisceau dirigé, on obtient une *gerbe* de photons. Cet aspect est précisément l'une des caractéristiques de l'émission *spontanée.*

échelles de temps, la plus petite étant très petite – mais évidemment jamais strictement nulle. En tout cas, ce dispositif vise à réaliser presque parfaitement ce que la Mécanique quantique appelle une mesure idéale, avec toutes ses conséquences y compris la réduction du paquet d'ondes, dans le cas où il y a deux résultats possibles : $|\mathrm{g}\rangle$ *ou* $|\mathrm{m}\rangle$. Un tel atome se prête donc *a priori* à l'observation de l'effet Zénon.

Itano et al. [98] ont d'une part réalisé cette expérience, d'autre part en ont donné la description théorique ; l'accord entre théorie et expérience est remarquable, justifiant que leur travail soit considéré comme la première mise en évidence de l'effet Zénon quantique, bien que leur interpétation a donné lieu à controverse (par exemple, voir [106] et [107]).

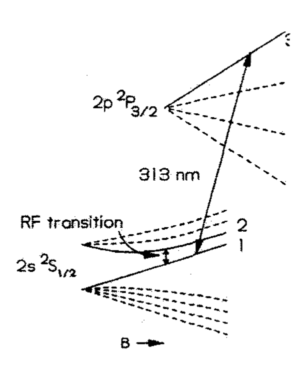


Figure 21.2: Diagramme des niveaux d'énergie de l'ion Be^+ dans un champ magnétique B (figure extraite de [98]). Observer que pour le couple d'états notés 1 et 2, on sort vite du régime Zeeman où $\Delta E \propto \mathcal{B}$, en raison d'un croisement de niveaux évité (voir problèmes 23.4.11 et 27.8.8).

L'expérience a été conduite sur une petite population d'ions[21] $^9\mathrm{Be}^+$ (environ 5000), en tout premier lieu soumise à un champ magnétique statique $\mathcal{B}$ assurant l'éclatement (*splitting*) Zeeman des états pertinents (voir fig. 21.3). Les états $|\mathrm{g}\rangle$ et $|\mathrm{m}\rangle$ sont deux sous-niveaux[22] hyperfins du fondamental $^2\mathrm{S}_{1/2}$ $((m_I, m_J) = (\frac{3}{2}, \frac{1}{2})$ et $(\frac{1}{2}, \frac{1}{2}))$. Un champ de radiofréquence[23] (RF) provoque la résonance de Rabi entre ces deux niveaux. Le rôle de $|\mathrm{e}\rangle$ est tenu par une composante appartenant au multiplet excité $^2\mathrm{P}_{3/2}$ $((m_I, m_J) = (\frac{3}{2}, \frac{3}{2}))$, couplée exclusivement au fondamental (l'émission spontanée de $^2\mathrm{S}_{1/2}$,$(\frac{1}{2}, \frac{1}{2})$ est totalement négligeable – voir le commentaire suivant (21.78)).

Soit $\hbar\Omega$ l'élément de matrice couplant les deux niveaux hyperfins entre lesquels

[21] L'ion Be^+ possède trois électrons et se comporte comme un alcalin, en l'occurrence le lithium Li.

[22] I se réfère au spin nucléaire, J au spin électronique. Sur la structure hyperfine, voir chapitre 27, section 27.4.

[23] Dans l'expérience, le champ magnétique vaut 0,819 T et la séparation de deux niveaux hyperfins est égale à 320,7 MHz ; l'amplitude du champ RF est ajustée de sorte que $\Omega T = \pi$ ("π-pulse"), avec $T = 256$ ms, soit $\Omega \simeq 4\pi \simeq 12$ rd/s ; la figure 21.3 schématise le diagramme d'énergie, pour une valeur donnée du champ magnétique dans la figure 21.3.

L'origine et la description de la structure hyperfine sont discutées dans le chapitre 27, section 27.4.

se produit l'oscillation de Rabi. On va voir que si le champ RF a une durée[24] T égale à $\frac{\pi}{\Omega}$ – on parle[25] alors de *π-pulse* –, l'ion partant du fondamental $|\mathrm{g}\rangle$ se retrouverait avec la probabilité $P_{\mathrm{m}} = 1$ dans l'état $|\mathrm{m}\rangle$ à l'issue de l'excitation RF s'il évoluait librement, en dehors de toute mesure. En fait, si une série de *pulses* optiques de durée très courte est simultanément envoyée sur l'ion, à des intervalles de temps réguliers $p\frac{T}{N}$ ($p = 1, 2, \ldots, N$), ces mesures successives vont ralentir la dynamique au point de donner $P_{\mathrm{m}} = 0$ (soit $P_{\mathrm{g}} = 1$) dans la limite d'un nombre infini de *pulses* : c'est très exactement l'effet Zénon quantique.

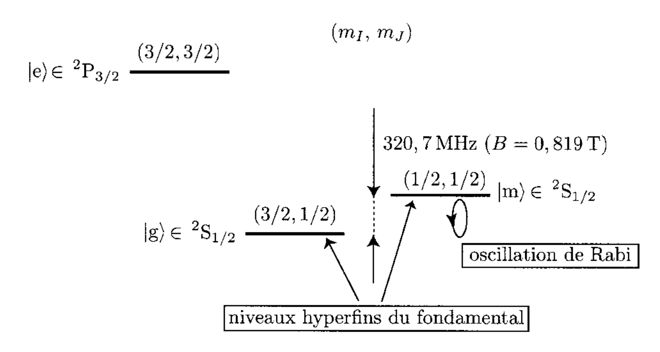


Figure 21.3: Atome à trois niveaux dans l'expérience de Itano et al. [98]. Chaque état est repéré par les valeurs du couple (m_I, m_J)

La description théorique repose fondamentalement sur l'idée suivante : entre deux *pulses*, le système évolue selon l'équation de Schrödinger (qui, de fait, n'implique que les deux niveaux hyperfins, ceux qui participent à l'oscillation de Rabi). Au contraire, l'effet d'un *pulse* (une "mesure") ne relève pas de cette équation et est décrit conformément au postulat de réduction du paquet d'ondes ; comme on vient de le voir, ceci revient à annuler *à la main* les éléments non-diagonaux de l'opérateur densité.

La dynamique s'effectue principalement de fait dans le sous-espace $(|\mathrm{g}\rangle, |\mathrm{m}\rangle)$, la réduction effective du paquet d'ondes conduisant soit vers $|\mathrm{m}\rangle$, soit vers $|\mathrm{g}\rangle$, suivant l'instant auquel arrive le *pulse* optique dans le cycle de Rabi. Le Hamiltonien $H(t)$ de l'atome couplé (semi-classiquement) au champ RF est ainsi représenté par une matrice 2×2, que l'on peut toujours prendre de trace nulle en choisissant le zéro d'énergie au milieu des énergies E_{g} et E_{m}. Si $\hbar\omega_0$ est la différence $E_{\mathrm{m}} - E_{\mathrm{g}}$, ω la pulsation du champ RF, la matrice de $H(t)$ sur la base ordonnée $(|\mathrm{m}\rangle, |\mathrm{g}\rangle)$ est de la forme ($\Omega \in \mathbb{R}$) :

$$H(t) = \frac{\hbar}{2} \begin{bmatrix} \omega_0 & \Omega \mathrm{e}^{-\mathrm{i}\omega t} \\ \Omega \mathrm{e}^{\mathrm{i}\omega t} & -\omega_0 \end{bmatrix} . \tag{21.49}$$

Comme on l'a vu à plusieurs reprises, on peut toujours, dans le cas d'un système à deux niveaux, représenter les matrices de trace nulle[26] en combinaison linéaire des trois

[24] Comme la pulsation RF ω est très grande devant Ω (fréquence de Rabi), cela a un sens de parler d'un champ RF de fréquence (presque) parfaitement déterminée de durée $T \sim \Omega^{-1} \gg \omega^{-1}$.

[25] Voir problème 18.7.11.

[26] Cette association a permis de définir l'homomorphisme entre SU(2) et SO(3) (voir problème 20.8.2).

matrices de Pauli ; pour des raisons évidentes, cette façon d'écrire s'appelle parfois *représentation en spin fictif* – voir p. 743 et le problème 18.7.11. C'est ainsi que, s'agissant de H, on peut écrire :

$$H(t) = \frac{\hbar}{2}\,\vec{\mathcal{H}}.\vec{\sigma} \ , \tag{21.50}$$

où $\vec{\mathcal{H}}$ a pour composantes les quantités $\hbar^{-1}\mathrm{Tr}(H\sigma_u)$, soit ici[27] :

$$\boxed{\vec{\mathcal{H}} = (\Omega\cos\omega t,\, \Omega\sin\omega t,\, \omega_0)} \tag{21.51}$$

Pour un système à deux niveaux, l'opérateur densité est aussi (représenté par) une matrice 2×2, d'éléments ρ_{gg}, ρ_{mm}, ρ_{mg} et ρ_{gm}, matrice hermitique si la base est orthonormée (ce qui est le cas). Comme de plus la trace de ρ est fixée et égale à 1, sa matrice possède trois degrés de liberté réels permettant à nouveau une représentation vectorielle[28] commode de ρ à l'aide d'un vecteur $\vec{R}$ de $\mathbb{R}^3$, dont les trois composantes réelles sont traditionnellement définies comme :

$$R_x = \rho_{\mathrm{gm}} + \rho_{\mathrm{mg}} \ , \qquad R_y = \mathrm{i}\,(\rho_{\mathrm{mg}} - \rho_{\mathrm{gm}}) \ , \qquad R_z = \rho_{\mathrm{mm}} - \rho_{\mathrm{gg}} \ . \tag{21.52}$$

Les cohérences de $\rho(t)$ sont associées aux composantes transverses R_x et R_y, alors que les populations sont contenues dans la composante longitudinale R_z. On vérifie sans peine que la relation précise entre ρ et $\vec{R}$ est :

$$\rho(t) = \frac{1}{2}\left(\mathbf{1}_2 + \vec{R}(t).\vec{\sigma}\right) \ , \qquad R_u = \mathrm{Tr}(\rho\sigma_u) \ ; \tag{21.53}$$

en particulier, on a :

$$\boxed{\rho_{\mathrm{mm}} = \frac{1}{2}(1 + R_z) \qquad \rho_{\mathrm{gg}} = \frac{1}{2}(1 - R_z)} \tag{21.54}$$

Avec cette paramétrisation, l'équation de Liouville (21.38) pour ρ s'écrit :

$$\mathrm{i}\hbar\,\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\hbar}{4}\,[\vec{\mathcal{H}}.\vec{\sigma},\, \vec{R}.\vec{\sigma}] \ , \tag{21.55}$$

et, après calcul[29], prend la forme gyroscopique :

$$\boxed{\frac{\mathrm{d}\vec{R}}{\mathrm{d}t} = \vec{\mathcal{H}}(t) \times \vec{R}} \tag{21.56}$$

La dépendance en temps du vecteur de rotation effectif $\vec{\mathcal{H}}(t)$ ne permet pas une résolution immédiate ; toutefois, ce vecteur tourne à vitesse angulaire constante, voir (21.51), ce

[27] Se souvenir de $\mathrm{Tr}(\sigma_u\sigma_v) = 2\delta_{uv}$, conséquence immédiate de $\sigma_u\sigma_v = \mathbf{1}_2\delta_{uv} + \mathrm{i}(1 - \delta_{uv})\varepsilon_{uvw}\sigma_w$.

[28] Quand il s'agit de l'opérateur densité, cette représentation est souvent appelée *représentation de Bloch*. Pour un état pur, $\vec{R}$ est un vecteur unitaire : en pareil cas, l'évolution de $\rho(t)$ se traduit par un mouvement de l'extrémité de $\vec{R}(t)$ sur la sphère unité, alors appelée *sphère de Bloch*. Toute perte de cohérence se traduit par une *contraction* de la longueur de $\vec{R}(t)$, voir plus loin...

[29] On utilise (18.143) au passage.

qui donne l'idée de passer dans le repère $Ox'y'z$ tournant à la pulsation ω autour de Oz, c'est-à-dire de poser[30] :

$$\rho'(t) \stackrel{\text{déf}}{=} U^\dagger(t)\rho(t)U(t) \ , \qquad U(t) \stackrel{\text{déf}}{=} \mathrm{e}^{-\mathrm{i}\,\frac{\omega t}{2}\sigma_z} \ ; \tag{21.57}$$

ceci fait, on obtient facilement l'équation satisfaite par le vecteur $\vec{R}'$ associé à ρ' :

$$\boxed{\frac{\mathrm{d}\vec{R}'}{\mathrm{d}t} = \vec{\mathcal{H}}' \times \vec{R}' \ , \qquad \vec{\mathcal{H}}' = (\Omega,\, 0,\, \omega_0 - \omega)} \tag{21.58}$$

techniquement, le point important est que $\vec{\mathcal{H}}'$ est maintenant indépendant du temps ; de ce fait l'intégration de (21.58) est immédiate. À la résonance stricte ($\omega = \omega_0$), $\vec{R}'$ précesse autour de Ox' à la pulsation Ω ; si à l'instant $t = 0$, l'ion est dans le fondamental $|\mathrm{g}\rangle$, alors :

$$\vec{R}'(t = 0) = (0,\, 0,\, -1) \ , \tag{21.59}$$

et – en l'absence de toute autre excitation – se retrouve[31] dans l'état $|\mathrm{m}\rangle$ au temps $T = \pi/\Omega$:

$$\vec{R}'(t = 0) = (0,\, 0,\, +1) \qquad \Longleftrightarrow \qquad \text{ion Be}^+ \text{ dans l'état } |\mathrm{m}\rangle \ ; \tag{21.60}$$

c'est bien parce que le vecteur $\vec{R}'$ a fait un demi-tour que l'on parle de π-*pulse* (pour le champ RF). À un instant quelconque, $\vec{R}'$ est quelque part dans le plan $y'Oz$ du repère tournant, les deux populations sont différentes de zéro, tout comme les cohérences : c'est l'oscillation de Rabi.

Introduisons maintenant le train régulier de *pulses*, comme dans l'expérience réalisée par Itano et al., chacun d'entre eux jouant le rôle d'une opération de mesure et survenant à des instants t_n calés sur la demi-période T, $t_n = n\,\frac{T}{N}$. Juste *avant* la première mesure, survenant à $t_1 = \frac{T}{N} = \frac{\pi}{N\Omega}$, on a :

$$\vec{R}'(t = \frac{\pi}{N\Omega}) = (0, \sin\Omega t, -\cos\Omega t)_{t=\frac{\pi}{N\Omega}} = (0,\, \sin\frac{\pi}{N},\, -\cos\frac{\pi}{N}) \ . \tag{21.61}$$

Survient alors le premier *pulse* optique ; traité comme une mesure idéale, il projette l'ion soit dans l'état $|\mathrm{m}\rangle$, soit dans l'état de $|\mathrm{g}\rangle$. Sur un ensemble d'ions représentés par l'opérateur densité, l'effet est d'annuler les éléments non-diagonaux (cohérences, composantes R_x et R_y, combinaisons linéaires de R'_x et R'_y) et de laisser intacts les éléments diagonaux (populations, composante R_z, égale à R'_z). À l'issue du premier *pulse* optique, l'opérateur densité est donc associé au vecteur :

$$\vec{R}'(t = \frac{\pi}{N\Omega} + 0) = (0,\, 0,\, -\cos\frac{\pi}{N}) \ . \tag{21.62}$$

La comparaison avec (21.59) montre que $\vec{R}'$ se retrouve dans la direction de départ mais a une longueur contractée d'un facteur $\cos\frac{\pi}{N}$. Chaque pulse successif donne lieu à une

[30] $U(t)$ est bien la rotation d'angle ωt autour de Oz, qui ne change pas la composante R_z : $R'_z = R_z$. La procédure suivie ici est la réplique de celle utilisée dans le problème 18.7.11.

[31] Ceci veut dire qu'une mesure effectuée à $t = \pi/\Omega + 0$ fournit le résultat : l'atome est en $|\mathrm{m}\rangle$ avec probabilité 1.

telle contraction : en vertu de la linéarité de l'équation gyroscopique (21.58), au bout de N mesures, le vecteur $\vec{R}'$ pointe dans la même direction qu'au départ, mais sa longueur est réduite d'un facteur $(\cos\frac{\pi}{N})^N$. Ainsi, à l'issue de la N^e mesure, on a :

$$\vec{R}'(t = T + 0) = \big(0,\, 0,\, -\cos^N \frac{\pi}{N}\big) \ . \tag{21.63}$$

En vertu de (21.52) et de $R'_z = R_z$, la probabilité pour un ion d'être dans $|\mathrm{m}\rangle$ à $T+0$ est :

$$P_\mathrm{m}(T+0) \ \equiv\ \rho_\mathrm{mm}(T+0) = \frac{1}{2}\big[1 + R_z(T+0)\big] = \frac{1}{2}\big(1 - \cos^N \frac{\pi}{N}\big) \ ; \tag{21.64}$$

dévéloppant le cosinus dans l'hypothèse $N \gg 1$, on a :

$$P_\mathrm{m}(T+0) \simeq \frac{1}{2}\big[1 - \mathrm{e}^{+N\,\ln(1-\frac{\pi^2}{2N^2})}\big] \simeq \frac{1}{2}\big(1 - \mathrm{e}^{-\frac{\pi^2}{2N}}\big) \simeq \frac{\pi^2}{4N} \ll 1 \ . \tag{21.65}$$

En l'absence de mesures répétées (de *pulses*), on trouverait $P_\mathrm{m}(T+0) = 1$, puisqu'il s'agit d'un π-*pulse* : (21.65) montre que la succession des mesures très rapprochées ruine la possibilité pour l'ion de filer dans l'état $|\mathrm{m}\rangle$ (plus précisément : d'y être trouvé). Inversement, la probabilité d'être dans l'état de départ $|\mathrm{g}\rangle$ est alors $P_\mathrm{g}(T+0) \simeq 1 - \frac{\pi^2}{4N} \sim 1$: au total, le train de mesures se succédant très rapidement empêche l'ion d'évoluer.

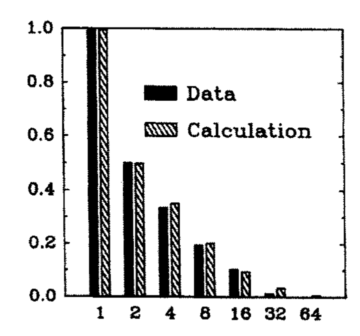


Figure 21.4: Variation de la probabilité pour un ion Be^+ d'être en $|\mathrm{m}\rangle$ en fonction du nombre N de pulses (figure extraite de [98]).

La décroissance de P_m avec le nombre de mesures (voir fig. 21.4) décrit quantitativement l'effet Zénon. La formule (21.65) donnée par Itano et al. se trouve être en accord remarquable avec les mesures, constituant ainsi la mise en évidence expérimentale de cet effet (voir [98] pour plus de détails) ; toutefois, comme déjà mentionné, cette interprétation a nourri une controverse dans la littérature (voir par exemple [106] et [107]).

21.2 Sauts quantiques

La notion de saut quantique remonte à la préhistoire de la Mécanique quantique. À la suite de l'introduction du photon par Einstein (effet photo-électrique, 1905), Bohr

suggéra en 1913 que l'interaction entre un atome et le rayonnement procède par absorption ou émission de photons, l'électron atomique effectuant des sauts d'une orbite à une autre (voir Tome I, section 6.2). À titre d'exemple, le processus d'émission suit la règle des fréquences de Bohr, qui n'est rien d'autre que l'expression de la conservation de l'énergie[32] :

$$E_\mathrm{i} = E_\mathrm{f} + h\nu \ . \tag{21.66}$$

E_i et E_f sont les énergies initiale et finale de l'atome, ν est la fréquence du photon reçu par le milieu extérieur à l'atome. En passant de l'état i à l'état f, l'atome effectue une transition, que l'on peut convenir d'appeler "saut quantique" (*"quantum jump"*).

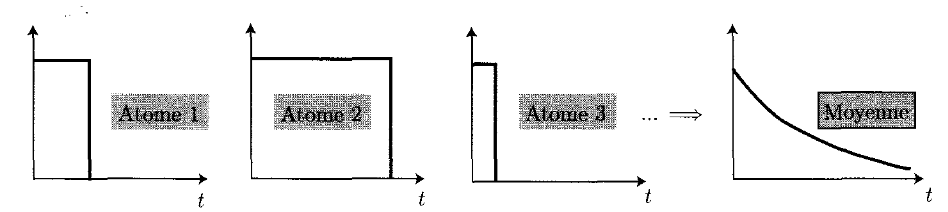


Figure 21.5: Émissions individuelles *vs* intensité moyenne de fluorescence.

Pendant des décennies, ce mécanisme élémentaire a été inobservable en tant que tel. En effet, une raie d'émission étant choisie, son observation s'est faite ususellement sur une vapeur atomique de basse densité contenant un très grand nombre d'atomes[33]. De ce fait, les transitions individuelles, noyées dans la masse et survenant à des instants aléatoires décorrélés statistiquement, construisent au total *de facto* une intensité de fluorescence macroscopique dont les fluctuations sont imperceptibles au point de confondre l'intensité aléatoire avec sa valeur *moyenne* (voir fig. 21.5). La mesure de cette intensité dans une vapeur où une fraction notable d'atomes est portée à l'état excité révèle un déclin essentiellement exponentiel, mettant en évidence une échelle de temps caractéristique $\tau_\mathrm{e} = \Gamma_\mathrm{e}^{-1}$, définissant la durée de vie *moyenne* d'un atome à l'état excité. En réalité, certains atomes se désexcitent bien avant τ_e, d'autres bien après et c'est seulement quand on fait la moyenne des temps de mort individuels que l'on retrouve la durée de vie moyenne[34].

[32]On suppose que l'atome est infiniment massif. Dans le cas contraire, il faut prendre en compte l'énergie de recul de l'atome (voir Tome I, problème 6.4.10).

[33]En 1952, Schrödinger écrivait :

> "... *we never experiment with just one electron or atom. In thought experiments we sometimes assume we do ; this invariably entails ridiculous consequences.*"

(cité en [108]). Comme l'écrit Harry Paul ([109], p. 36), commentant les expériences de Sauter et al. [110] :

> "*Schrödinger's scepticism could not have been disproved more convincingly !*"

[34]Il en va exactement de même pour la décroissance radioactive. Au début (beaucoup de noyaux actifs), l'énormité du nombre de noyaux présents dans une source macroscopique engendre un déclin exponentiel ne présentant aucune anomalie, aucune marche n'étant visible à la sensibilité affichée (pendant l'intervalle élémentaire de comptage, δt, il se produit un nombre faramineux de désintégrations) ; pendant cette première phase, il est licite de confondre la valeur de la variable aléatoire N (nombre de noyaux actifs à l'instant considéré, observée) avec sa valeur *moyenne*. Au contraire, au bout d'un temps très long

En termes probabilistes, la situation décrite ci-dessus peut se modéliser simplement comme suit. L'atome i se désexcite à l'instant τ_i, réputé aléatoire ; écartant pour des raisons physiques évidentes[35] tout phénomène coopératif entre atomes, les différents instants τ_j sont statistiquement indépendants. Le nombre d'atomes $N(t)$ encore excités à l'instant t est donné par la fonction indicatrice :

$$N(t) = \sum_{i=1}^{N_0} \theta(\tau_i - t) \ , \tag{21.67}$$

où $\theta(x)$ est la fonction échelon-unité ($\theta(x<0) = 0$, $\theta(x>0) = 1$). La proportion d'atomes encore excités à cet instant est $N(t)/N_0$; pourvu que $N_0 \ggg 1$, cette fraction se confond pratiquement avec la moyenne sur l'instant τ, calculable avec la loi de distribution de τ, $\rho(\tau)$:

$$\frac{N(t)}{N_0} = \frac{1}{N_0}\sum_{i=1}^{N_0} \theta(\tau_i - t) \rightarrow \langle\theta(\tau - t)\rangle \stackrel{\text{déf}}{=} \int_0^{+\infty} \theta(\tau - t)\rho(\tau)\,\mathrm{d}\tau \ . \tag{21.68}$$

On veut que cette moyenne soit l'exponentielle $\mathrm{e}^{-\Gamma_\mathrm{e} t}$, ce qui va contraindre la loi $\rho(\tau)$, et en fait permettre de la trouver sans ambiguïté. Pour cela, on utilise la transformation de Laplace ; l'image de Laplace de $\theta(\tau - t)$ est $\int_0^\tau \mathrm{e}^{-zt}\,\mathrm{d}\tau = z^{-1}(1 - \mathrm{e}^{-z\tau})$, d'où, par la formule d'inversion de Laplace, puis en prenant la moyenne membre à membre :

$$\theta(\tau - t) = \frac{1}{2\mathrm{i}\pi}\int_{\mathcal{B}} \mathrm{e}^{zt}\frac{1-\mathrm{e}^{-z\tau}}{z}\,\mathrm{d}z \quad \Longrightarrow \quad \langle\theta(\tau - t)\rangle = \frac{1}{2\mathrm{i}\pi}\int_{\mathcal{B}} \mathrm{e}^{zt}\frac{1-\langle\mathrm{e}^{-z\tau}\rangle}{z}\,\mathrm{d}z \ , \tag{21.69}$$

$\mathcal{B}$ désignant la droite de Bromwich. Pour que l'intégrale de droite reproduise $\mathrm{e}^{-\Gamma_\mathrm{e} t}$, il faut que le facteur de e^{zt} dans l'intégrand soit égal à $(z+\Gamma_\mathrm{e})^{-1}$, ce qui donne l'expression pour la moyenne cherchée : $\langle \mathrm{e}^{-z\tau}\rangle = \Gamma_\mathrm{e}/(z+\Gamma_\mathrm{e})$; on en déduit :

$$\int_0^{+\infty} \mathrm{e}^{-z\tau}\rho(\tau)\,\mathrm{d}\tau = \frac{\Gamma_\mathrm{e}}{z+\Gamma_\mathrm{e}} \quad \Longleftrightarrow \quad \rho(\tau) = \Gamma_\mathrm{e}\mathrm{e}^{-\Gamma_\mathrm{e}\tau} \ . \tag{21.70}$$

L'instant τ de désexcitation aléatoire d'un atome est donc distribué suivant une loi exponentielle, paramétrée précisément par la durée de vie moyenne observable de l'état excité Γ_e. C'est ainsi que la moyenne statistique des $\theta(\tau_i - t)$ peut reconstituer la fraction excitée à déclin exponentiel que l'on observe avec une vapeur atomique diluée, dans la limite bien sûr des très grands nombres d'atomes N_0.

Ainsi, pour espérer voir les sauts quantiques, il faut travailler avec un petit nombre de sources (atomes, ions), mais encore faut-il que le signal attestant de la présence de

devant la période radioactive, des fluctuations toujours présentes deviennent visibles (puisqu'il reste alors un petit nombre de noyaux encore excités) et l'exponentielle, mesurée pour une source donnée, devient "bruitée". Analysant le déclin radioactif comme il doit l'être (un processus *aléatoire*), on trouve d'ailleurs que les fluctuations relatives $\frac{\Delta N}{\langle N\rangle}$ *divergent* quand $t \rightarrow +\infty$ (voir Tome I, p. 44).

[35]On est loin du régime de *superradiance* sommairement décrit dans le Tome I p. 171. Le calcul qui suit requiert l'indépendance statistique des τ_i puisqu'il repose sur l'existence d'une loi de probabilité $\rho(\tau)$ pour la durée de vie d'*un* atome, vivant sa vie en solitaire, indépendamment des autres.

quelques atomes dans un état ou un autre soit assez intense pour pouvoir être observé et mesuré. Assez récemment (au début des années 1980), il est devenu possible de piéger quelques ions et de les refroidir à l'aide de lasers, autorisant leur observation dans des conditions quasi-idéales, rendant réalisable une expérience proposée antérieurement par Dehmelt [111] ; il s'agit d'une expérience de double résonance optique dont le principe est le suivant.

Soit un atome à trois niveaux, comme celui introduit dans la section 21.1 (voir fig. 21.6). Les deux niveaux $|\mathrm{m}\rangle$ et $|\mathrm{e}\rangle$ sont supposés avoir des énergies voisines, les deux différences $E_{\mathrm{m}} - E_{\mathrm{g}}$ et $E_{\mathrm{e}} - E_{\mathrm{g}}$ sont dans le domaine optique ou UV. Typiquement, la durée de vie de $|\mathrm{e}\rangle$ est de l'ordre de 10^{-8} s, celle de $|\mathrm{m}\rangle$ est de l'ordre de la seconde. Ces deux ordres de grandeurs, très différents, indiquent que les probabilités des transitions sont elles-mêmes très différentes : la transition $|\mathrm{g}\rangle \leftrightarrow |\mathrm{m}\rangle$ est très rare[36], comparée à $|\mathrm{g}\rangle \leftrightarrow |\mathrm{e}\rangle$. Par ailleurs, l'atome est soumis à deux lasers ; le premier, très intense, sature la transition $|\mathrm{g}\rangle \leftrightarrow |\mathrm{e}\rangle$ – c'est-à-dire égalise très vite les populations des états $|\mathrm{g}\rangle$ et $|\mathrm{e}\rangle$ –, tandis que le second est calé sur la transition rare $|\mathrm{g}\rangle \leftrightarrow |\mathrm{m}\rangle$.

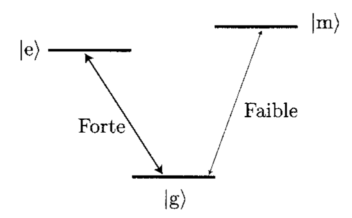


Figure 21.6: Atome à trois niveaux pour l'expérience de double résonance optique.

Si l'état de l'ion est le fondamental $|\mathrm{g}\rangle$ ou l'état excité $|\mathrm{e}\rangle$, le laser intense induit un mouvement de va-et-vient rapide entre ces deux états ; cette "oscillation" est rapide puisque la transition est permise et que la durée de vie de $|\mathrm{e}\rangle$ est courte ; elle produit une fluorescence intense, facilement observable. Cependant, l'autre laser peut à un certain instant porter l'atome à l'état $|\mathrm{m}\rangle$; il faut du temps pour cela, puisque la transition est rare – mais elle finit bien par se produire. Dès lors, la fluorescence s'arrête et reste "noire" tant que l'atome n'est pas redescendu en $|\mathrm{g}\rangle$ à partir de $|\mathrm{m}\rangle$. Quand ceci se produit, la fluorescence recommence, et ainsi de suite. Au total, dans l'expérience proposée par Dehmelt, on doit assister à une fluorescence intermittente[37] illustrée schématiquement

[36] L'aller-retour entre $|\mathrm{g}\rangle$ et $|\mathrm{m}\rangle$ n'est pas une oscillation du type Rabi : c'est simplement le mécanisme ordinaire d'absorption et d'émission (essentiellement spontanée) entre deux états couplés par un champ électromagnétique. Dans le domaine des micro-ondes, l'émission induite l'emporte sur l'émission spontanée, c'est le contraire dans le domaine optique ou UV. Cette règle générale possède des exceptions : même dans le domaine optique, l'émission spontanée peut être faible en comparaison de l'émission induite pourvu que l'on utilise un laser très intense (alors, on peut avoir $A \ll BW$, dans les notations introduites peu après).

[37] La transition faible survient à des instants aléatoires ; si ces derniers sont décorrélés, ce processus aléatoire binaire (prenant deux valeurs) est connu sous le nom de *processus des télégraphistes*. Le mot *intermittence* n'est pas à prendre ici dans une acception récente qui qualifie la dynamique *par bouffées* de certains systèmes non-linéaires [112].

sur la figure 21.7. Par ailleurs, on peut affirmer que quand la fluorescence est absente, l'ion est dans l'état $|\mathrm{m}\rangle$ et que quand elle se produit, il est dans le sous-espace (secteur) $(|\mathrm{g}\rangle, |\mathrm{e}\rangle)$. En un sens, l'intermittence de la fluorescence est une succession de réductions de paquets d'ondes, ou, dit autrement, l'observation de l'état de fluorescence constitue une observation directe des sauts quantiques[38] impliquant l'état $|\mathrm{m}\rangle$: l'extinction de la fluorescence signifie que l'ion vient d'arriver en $|\mathrm{m}\rangle$, sa réapparition que l'ion vient de regagner le secteur $(|\mathrm{g}\rangle, |\mathrm{e}\rangle)$.

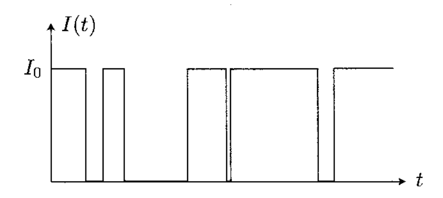


Figure 21.7: Fluorescence intermittente d'un atome unique (comparer avec fig. 21.9).

Le premier traitement théorique a été donné par Cook et Kimble [113] et repose sur des équations *cinétiques* mettant en jeu les coefficients d'Einstein, qui font l'objet de l'aparté suivant.

■ Les coefficients d'Einstein

Les *coefficients d'Einstein* A_ν et B_ν ont été introduits par Einstein en 1917 pour décrire quantitativement les échanges d'énergie entre rayonnement et matière [114]. Dans l'hypothèse où les atomes sont des systèmes dont les états ont une énergie *quantifiée*, chaque atome peut prendre de l'énergie au champ en passant d'un état d'énergie E_m à un état d'énergie $E_n > E_m$ (absorption), ou au contraire céder de l'énergie au champ en effectuant une transition $E_n \rightarrow E_m$ (émission) ; dans tous les cas, suivant la règle de Bohr, la fréquence $\nu_{nm} \stackrel{\text{déf}}{=} h^{-1}(E_n - E_m)$ est celle de la lumière émise ou absorbée. Einstein introduit plus précisément *trois* processus élémentaires : l'absorption, l'émission *induite* (ou *stimulée*) et l'émission *spontanée* et, pour chacun d'entre eux, définit une certaine "constante de réaction" permettant d'écrire des équations cinétiques d'évolution pour les populations atomiques, $N_n(t)$ étant le nombre[39] d'atomes se trouvant à l'instant t dans l'état d'énergie E_n. Dans tous les cas, la contribution d'un processus donné à toute variation $\mathrm{d}N$ de population est supposée proportionnelle au nombre d'atomes se trouvant dans l'état qui est le point de *départ* de la transition.

[38] Le signal observé est la fluorescence induite par le premier laser, dont la puissance spectrale détermine l'intensité de la fluorescence ; la question n'est pas de pouvoir observer le *photon unique* émis lors de la transition $|\mathrm{m}\rangle \rightarrow |\mathrm{g}\rangle$.

[39] Plus précisément, N_n est la *moyenne* du nombre d'atomes dans l'état d'énergie E_n, d'ailleurs, l'état d'équilibre sera défini plus bas comme celui où toutes les dérivées $\dot{N}_n$ sont nulles, alors que même dans l'état d'équilibre, il y a des *fluctuations*. À nouveau, la confusion entre la valeur instantanée d'une variable aléatoire et sa valeur moyenne est sans incidence quand il s'agit de grands nombres.

En ce qui concerne l'absorption et l'émission induite, Einstein postule que la variation de population qui en résulte est proportionnelle à l'intensité $W(\nu_{nm})$ du champ électromagnétique ($W(\nu_{nm}) = W(\nu_{mn})$) à la fréquence de la transition $E_n \leftrightarrow E_m$; le coefficient de proportionnalité correspondant est noté B, plus précisément B_{nm} quand il s'agit de la transition $E_m \rightarrow E_n$ (B est pris positif). Bien évidemment, le nombre total de transitions est proportionnel au nombre d'atomes situés dans l'état de départ ; ainsi, pour un couple d'états (n, m) donné, la variation δN_n entre t et $t + \delta t$ résultant de ces deux seuls processus est égale à $(-B_{mn}N_n + B_{nm}N_m)W(\nu_{nm})\delta t$.

Pour l'émission spontanée – nouveau concept introduit par Einstein –, mais dont l'existence est prouvée par le seul fait qu'un atome excité finit, *en l'absence de champ*, par retomber dans l'état fondamental, Einstein introduit un coefficient inédit A (également positif), tel que la désexcitation d'un atome par émission spontanée vers un état d'énergie inférieure E_m donne lieu à la variation $\delta N_n = -A_{mn}N_n\delta t$ entre t et $t + \delta t$. Bien sûr, aucune intensité de champ n'apparaît dans cette variation élémentaire[40].

Prenant en compte l'ensemble des processus possibles pour tous les couples d'états, on peut ainsi écrire un système d'équations cinétiques qui font le bilan de toutes les réactions ; en passant à la limite $\delta t \rightarrow 0$, on obtient ainsi un système différentiel :

$$\frac{\mathrm{d}N_n}{\mathrm{d}t} = -\sum_{m\neq n} B_{mn}W(\nu_{nm})N_n + \sum_{m\neq n} B_{nm}W(\nu_{nm})N_m - \sum_{m<n} A_{mn}N_n + \sum_{m>n} A_{nm}N_m \ , \tag{21.71}$$

avec une équation de ce type pour chaque niveau atomique ; la notation $\sum_{m<n}$ signifie que l'on somme sur les états d'énergie E_m *inférieure* à E_n, et inversement. Le premier terme au second membre est la dépopulation de l'état n par les transitions induites de n vers tous les autres états m ; le second est le processus inverse (repopulation de l'état n), toujours induit par le champ. Le troisième est la perte de l'état n par émission spontanée vers les états m d'énergie $E_m < E_n$; le dernier terme est la repopulation de l'état n par désexcitation spontanée des états plus fortement excités. On montre ci-dessous que, en l'absence de dégénérescence, $B_{nm} = B_{mn}$ (voir (21.78)) ; en vertu de cette symétrie et une fois rassemblés, les deux termes induits s'écrivent $-\sum_{m\neq n} B_{mn}W(\nu_{nm})(N_n - N_m)$, expression qui a la forme d'une loi du genre Fick, $B_{mn}W(\nu_{nm})$ jouant le rôle d'une *vitesse*, la différence $(N_n - N_m)$ celui d'un gradient (discret) de concentration. Une telle contribution a donc le sens d'un *courant*, qui est induit par le champ, courant qui est nul si les deux populations sont égales (transition saturée).

Le système (21.71) est un exemple d'*équation-maîtresse*, un type d'équation de nature universelle[41] puisque, moyennant des hypothèses précises permettant de transcrire algébriquement les processus élémentaires, elle ne fait qu'exprimer un *bilan* dans

---

[40] Quantiquement, ceci signifie que l'émission spontanée se produit quand le champ est nul *en moyenne*. Elle résulte en fait des *fluctuations* du champ qui, comme tout objet quantique, a une nature foncièrement aléatoire : quand le champ est nul en moyenne, c'est parce qu'il est dans son état (quantique) fondamental. Dans cet état *vide* de tout photon (en moyenne) subsistent néanmoins des fluctuations (quantiques) irréductibles ; de la même façon, un simple oscillateur harmonique dans son état fondamental présente des fluctuations Δx et Δp, dont le produit est verrouillé inférieurement par la constante de Planck.

[41] On trouve ce type d'équations en Chimie, lorsqu'il s'agit de traduire l'évolution des concentrations de plusieurs espèces chimiques se transformant les unes dans les autres ; ce cadre est d'ailleurs un paradigme

un système fermé. La somme de toutes les équations membre à membre donne zéro, traduction de la conservation de la somme $\sum_n N_n$.

Le système (21.71) est supposé valide à tout instant, y compris pour décrire une évolution hors d'équilibre ; bien évidemment, il doit aussi restituer, à $t = +\infty$, la situation connue par ailleurs de l'équilibre thermodynamique, quand le rayonnement et la matière sont en présence : on doit donc retrouver en particulier l'état d'équilibre du rayonnement à la température T, c'est-à-dire... la loi de Planck (voir Tome I, section 4.2) – c'était d'ailleurs l'un des objectifs d'Einstein. L'état d'équilibre est caractérisé par le fait que tous les premiers membres du système (21.71) sont nuls, d'où :

$$0 = \sum_{m \neq n} (-B_{mn} N_n^{\text{eq}} + B_{nm} N_m^{\text{eq}}) W(\nu_{nm}) - \sum_{m<n} A_{mn} N_n^{\text{eq}} + \sum_{m>n} A_{nm} N_m^{\text{eq}} \ . \qquad (21.72)$$

Telle qu'il est écrit, ce système de relations n'est guère utilisable ; par une argumentation subtile et difficile, van Kampen [116] a montré qu'en fait, en l'absence de champ magnétique (schématiquement parlant), une condition *suffisante* autrement plus simple est satisfaite à l'équilibre, exprimant que le bilan est nul *pour un couple d'états donné*. Raisonnant avec le couple (n, m) avec $E_n > E_m$, cette condition *suffisante* s'écrit :

$$0 = (-B_{mn} N_n^{\text{eq}} + B_{nm} N_m^{\text{eq}}) W(\nu_{nm}) - A_{mn} N_n^{\text{eq}} \ . \qquad (21.73)$$

Ces conditions suffisantes portent le nom de *bilan détaillé* (*detailed balance*) et assurent que l'équilibre est atteint[42]. Cela étant admis, l'équilibre canonique à la température T permet d'écrire[43] :

$$\frac{N_n^{\text{eq}}}{N_m^{\text{eq}}} = \mathrm{e}^{-\beta(E_n - E_m)} \ ; \qquad (21.75)$$

par substitution dans la condition de bilan détaillé (21.73), on obtient une première relation entre les coefficients d'Einstein :

$$0 = (-B_{mn} \mathrm{e}^{-\beta h \nu_{nm}} + B_{nm}) W(\nu_{nm}) - A_{mn} \mathrm{e}^{-\beta h \nu_{nm}} \ . \qquad (21.76)$$

Par ailleurs, on peut prendre pour mesurer l'intensité du rayonnement soit sa densité volumique d'énergie u, soit sa luminance L (voir Tome I, sous-section 4.1.2), les deux quantités étant reliées l'une à l'autre par $4\pi L = cu$ (Tome I, éq. (4.19)), ce qui conduit à

---

pour l'occurrence de comportements exotiques puisque toute réaction bimoléculaire se traduit par des *non-linéarités*. Un exemple célèbre – qui a valu bien du souci à ses inventeurs car on ne croyait pas à l'époque qu'une telle réaction fut possible – est la réaction *oscillante* de Belouzov - Zhabotinskii [115].

C'est encore une équation de ce type qui, moyennant des hypothèses supplémentaires concernant la nature stochastique des processus, conduit à la célèbre équation de Fokker - Planck, dont la version la plus simple est l'équation classique de la diffusion. Enfin, la relation fondamentale pour les processus markoviens (équation de Bachelier - Chapman - Kolmogorov) est encore un avatar d'équation-maîtresse.

[42] Dans des présentations plus élémentaires, on considère d'emblée *un* seul couple de niveaux pour les atomes ; par construction, on court-circuite alors la nécessité d'invoquer le principe du bilan détaillé.

[43] En toute rigueur, il faudrait introduire la dégénérescence g_n de l'état n, et écrire plus précisément :

$$\frac{N_n^{\text{eq}}}{N_m^{\text{eq}}} = \frac{g_n}{g_m} \mathrm{e}^{-\beta(E_n - E_m)} \ . \qquad (21.74)$$

Dans un but de simplicité, on ignore cette complication inessentielle dans la suite, qui donne juste un facteur g_m/g_n en plus aux second membres de (21.78).

deux définitions légèrement différentes des coefficients A et B ; optant pour u, $W(\nu)$ n'est autre que la densité prévue par la loi de Planck, soit $W(\nu) = (8\pi h\nu^3/c^3)\,(\mathrm{e}^{\beta h\nu} - 1)^{-1}$ (Tome I, éq. (4.99)) ; dès lors, (21.76) devient :

$$0 = \frac{8\pi h\nu_{nm}^3}{c^3}\frac{1}{\mathrm{e}^{\beta h\nu_{nm}} - 1}(-B_{mn}\mathrm{e}^{-\beta h\nu_{nm}} + B_{nm}) - A_{mn}\mathrm{e}^{-\beta h\nu_{nm}} \ . \tag{21.77}$$

Ceci devant être vrai à toute température, l'identification après rassemblement des termes constants et de ceux ne dépendant que de $\mathrm{e}^{\beta h\nu_{nm}}$ donne les deux relations :

$$\boxed{\frac{A_{mn}}{B_{nm}} = \frac{8\pi h\nu_{nm}^3}{c^3} \simeq 6,2\times 10^{-31}(\nu_{\mathrm{GHz}}^3)\,\mathrm{J\,s\,m^{-3}} \qquad \frac{B_{mn}}{B_{nm}} = 1} \tag{21.78}$$

La dépendance en ν^3 permet de comprendre pourquoi l'émission spontanée est négligeable à basse fréquence, ce qui justifie de l'oublier quand il est question de l'oscillation de Rabi.

L'existence de A_{nm}, avérée par l'émission spontanée, est visiblement liée au fait que les états excités atomiques sont, à l'équilibre thermique, forcément moins peuplés que les états de plus basse énergie (ce qu'expriment les facteurs de Boltzmann $\mathrm{e}^{-\beta E_n}$) ; la symétrie entre émission et absorption induite entraîne que, s'il n'était d'émission spontanée et à l'équilibre, les populations des états seraient *égales* entre elles. Sur un tout autre plan, la traduction formelle de l'existence de l'émission spontanée est plutôt technique : elle réside dans la non-commutation des opérateurs de champ dans la description quantique du champ électromagnétique (voir chapitre 25, notamment le commentaire suivant (25.67)). ■

Une fois introduits les coefficients d'Einstein, il est possible d'exposer le premier traitement théorique du signal intermittent de fluorescence – qui signe les *sauts quantiques* de Bohr –, dû à Cook et Kimble [113], et qui repose précisément sur les équations cinétiques d'Einstein pour les populations P_{a} (a = g, m, e) des états atomiques (conformément à l'usage, on désigne maintenant par P_n ce qui est noté N_n dans le traitement historique d'Einstein) ; par la suite, Zoller et al. [117] ont donné un traitement microscopique du phénomène, fondé sur un traitement quantique complet où le rayonnement est également quantifié.

En notant simplement les coefficients d'Einstein A_{a} ($\equiv A_{\mathrm{ga}}$) et $B_{\mathrm{bc}} = B_{\mathrm{cb}}$, et W_i les densités spectrales des deux lasers, les équations cinétiques s'écrivent :

$$\dot{P}_{\mathrm{g}} = -(B_{\mathrm{eg}}W_1 + B_{\mathrm{mg}}W_2)P_{\mathrm{g}} + (A_{\mathrm{e}} + B_{\mathrm{eg}}W_1)P_{\mathrm{e}} + (A_{\mathrm{m}} + B_{\mathrm{mg}}W_2)P_{\mathrm{m}} \ , \tag{21.79}$$

$$\dot{P}_{\mathrm{e}} = -(A_{\mathrm{e}} + B_{\mathrm{eg}}W_1)P_{\mathrm{e}} + B_{\mathrm{eg}}W_1P_{\mathrm{g}} \ , \tag{21.80}$$

$$\dot{P}_{\mathrm{m}} = -(A_{\mathrm{m}} + B_{\mathrm{mg}}W_2)P_{\mathrm{m}} + B_{\mathrm{mg}}W_2P_{\mathrm{g}} \ , \tag{21.81}$$

ce que l'on peut traduire matriciellement par $|\dot{P}\rangle = -\mathbf{\Gamma}|P\rangle$. On vérifie que $\sum \dot{P}_{\mathrm{a}} = 0$, relation qui exprime la conservation de la somme $P_{\mathrm{g}} + P_{\mathrm{e}} + P_{\mathrm{m}}$, égale à 1 à tout temps (la somme des éléments d'une colonne de $\mathbf{\Gamma}$ est nulle, d'où une valeur propre nulle pour

$\mathbf{\Gamma}$) ; par hypothèse, $A_\mathrm{m} \ll A_\mathrm{e}$. Le système (21.79) - (21.81) a une solution d'équilibre $|P^\mathrm{eq}\rangle$ obtenue en annulant tous les premiers membres ; on trouve ainsi :

$$\frac{P_\mathrm{e}^\mathrm{eq}}{P_\mathrm{g}^\mathrm{eq}} = \frac{B_\mathrm{eg}W_1}{A_\mathrm{e} + B_\mathrm{eg}W_1} \ , \qquad \frac{P_\mathrm{m}^\mathrm{eq}}{P_\mathrm{g}^\mathrm{eq}} = \frac{B_\mathrm{mg}W_2}{A_\mathrm{m} + B_\mathrm{mg}W_2} \ ; \tag{21.82}$$

si on peut négliger le terme d'émission spontanée dans chaque dénominateur, les populations d'équilibre sont égales entre elles et valent chacune 1/3. Cette situation est atteinte longtemps après la plus longue échelle de temps, usuellement A_m^{-1}, qui est gigantesque par rapport à toutes les autres.

Avec W_2 strictement nul (deuxième laser coupé) et $P_\mathrm{g}(t = 0) = 1$, on obtient la quasi-égalité[44] des deux probabilités P_g et P_e ($P_\mathrm{g} \simeq P_\mathrm{e} = 1/2$) après un temps de l'ordre de $(A_\mathrm{e} + 2B_\mathrm{eg}W_1)^{-1}$: on dit alors que la transition (forte) e $\leftrightarrow$ g est *saturée* (les populations des deux états impliqués sont égales entre elles). Avec W_2 non nul (mais $B_\mathrm{mg}W_2 \ll B_\mathrm{eg}W_1$) la transition forte reste saturée – ce que l'on obtient quand W_1 est très grand (assurant que $A_\mathrm{e} \ll B_\mathrm{eg}W_1$) – les populations P_e et P_g étant à peu près égales[45] ; on se retrouve de fait avec une sorte de système à deux niveaux : l'atome est soit dans le secteur ($|\mathrm{g}\rangle$, $|\mathrm{e}\rangle$) avec des populations presque égales $P_\mathrm{g} \simeq P_\mathrm{e}$, soit dans l'état $|\mathrm{m}\rangle$; il suffit dès lors de déterminer les deux populations $\mathcal{P}_\pm$:

$$\mathcal{P}_+ = \mathcal{P}_\mathrm{m} \qquad \mathcal{P}_- = \mathcal{P}_\mathrm{g} + \mathcal{P}_\mathrm{e} \ , \tag{21.84}$$

sachant que, dans ce régime, on a :

$$P_\mathrm{g} \simeq P_\mathrm{e} \simeq \frac{1}{2}\mathcal{P}_- \ . \tag{21.85}$$

Reportant les nouvelles variables $\mathcal{P}_\pm$ dans (21.81), il vient compte tenu de (21.85) :

$$\dot{\mathcal{P}}_+ = -(A_\mathrm{m} + B_\mathrm{mg}W_2)\mathcal{P}_+ + \frac{1}{2}B_\mathrm{mg}W_2\,\mathcal{P}_- \ , \tag{21.86}$$

soit :

$$\boxed{\dot{\mathcal{P}}_+ = -R_-\mathcal{P}_+ + R_+\mathcal{P}_-} \tag{21.87}$$

[44]Ceci provient de $A_\mathrm{e} \ll B_\mathrm{eg}W_1$, réalisable avec un W_1 suffisamment grand. Quand $W_2 = 0$ et avec $P_\mathrm{g}(t = 0) = 1$, la solution exacte de (21.79) - (21.80) est :

$$P_\mathrm{g}(t) = \frac{1}{A_\mathrm{e} + 2B_\mathrm{eg}W_1}\left[A_\mathrm{e} + B_\mathrm{eg}W_1 - B_\mathrm{eg}W_1\mathrm{e}^{-(A_\mathrm{e}+2B_\mathrm{eg}W_1)t}\right] \ , \quad P_\mathrm{e}(t) = 1 - P_\mathrm{g}(t) \ . \tag{21.83}$$

En l'absence d'émission spontanée ($A_\mathrm{e} = 0$), on aurait $P_\mathrm{g}(+\infty) = P_\mathrm{e}(+\infty) = 1/2$: la transition g $\leftrightarrow$ e serait alors strictement saturée.

[45]Les deux populations P_g et P_e sont presque égales, mais diffèrent de 1/2. Sur une échelle de temps courte par rapport à $(B_\mathrm{mg}W_2)^{-1}$, tout se passe comme si le niveau métastable était absent et ces deux populations valent à peu près 1/2. En revanche, aux temps longs, bien après que le niveau $|\mathrm{m}\rangle$ a eu le temps de jouer son rôle et toujours dans l'hypothèse où la transition forte reste saturée, toutes les populations tendent vers 1/3 quand l'état $|\mathrm{m}\rangle$ a une très longue durée de vie (pour la discussion détaillée, voir [108]). En définitive, avec les hypothèses en vigueur, le système se comporte d'abord comme un atome à deux niveaux, puis comme un atome à trois niveaux.

Ces éléments semi-quantitatifs sont confirmés par la résolution du système ((21.79) - (21.81)) dans ce régime de saturation.

où $R_+ = \frac{1}{2}B_{\rm mg}W_2$, $R_- = A_{\rm m} + B_{\rm mg}W_2$. Comme $\mathcal{P}_+ + \mathcal{P}_- = 1$, on en déduit :

$$\boxed{\dot{\mathcal{P}}_- = -R_+\mathcal{P}_- + R_-\mathcal{P}_+} \tag{21.88}$$

L'inspection de ces équations révèle que R_- allume la fluorescence, alors que R_+ l'éteint, ce que schématise la figure 21.8.

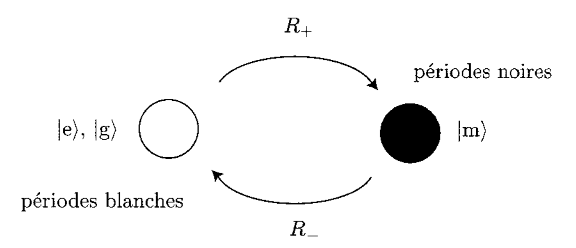


Figure 21.8: R_+ éteint la fluorescence, R_- la rallume.

L'état d'équilibre des équations (21.87) et (21.88) est :

$$\mathcal{P}_{+\,\rm eq} = \frac{R_+}{R_+ + R_-}\ , \qquad \mathcal{P}_{-\,\rm eq} = \frac{R_-}{R_+ + R_-}\ . \tag{21.89}$$

Ces probabilités donnent la *moyenne* temporelle du signal de fluorescence, une fois les transitoires éteints ; si I_0 désigne l'intensité (constante) d'une période allumée (blanche) :

$$\bar{I} = \mathcal{P}_{+\,\rm eq} \times 0 + \mathcal{P}_{-\,\rm eq} \times I_0 = I_0\mathcal{P}_{-\,\rm eq} = \frac{R_-}{R_+ + R_-}\, I_0 \tag{21.90}$$

En revanche, les probabilités stationnaires (21.89) ne disent évidemment rien sur les *fluctuations* de ce signal. Celles-ci ne peuvent être obtenues que comme sous-produits d'une description plus fine du processus stochastique, que l'on suppose stationnaire (toutes les probabilités à deux temps ne dépendent que de la différence des temps, toutes les probabilités à un temps sont indépendantes du temps).

Désignons par $P_{n\,\pm}(t)$ la probabilité qu'il y ait entre les instants t_0 et $t_0 + t$, n transitions d'un secteur à l'autre *et* que l'état final (à l'instant t) soit $\pm$; par construction, les valeurs de ces probabilités à $t = 0$ sont 0 et 1, suivant l'état à $t = 0$ (aucune transition ne peut se produire sur un intervalle de temps nul). Par l'hypothèse de stationnarité (invariance par translation dans le temps, indifférence de l'instant appelé t_0), ces probabilités ne dépendent que de t. Pour obtenir leur équation d'évolution, il suffit à nouveau de faire un bilan classique de probabilité entre deux instants très proches t et $t + \Delta t$, et de prendre la limite $\Delta t \to 0$. Par exemple, entre t et $t + \Delta t$, $P_{n\,+}$ peut, par un seul *flip* au plus, évoluer de deux façons :

1. il ne se passe rien, probabilité $1 - R_-\Delta t$

2. il y a eu $n-1$ transitions jusqu'à t, qui ont conduit le système dans l'état $-$ et l'une se produit entre t et $t+\Delta t$, probabilité $R_+\Delta t$

Évidemment, bien d'autres processus sont possibles, mais comme ils impliquent plus d'un *flip*, leurs probabilités sont de ce fait en $\Delta t^{k\geq 2}$: ils ne contribuent donc pas, dans la limite $\Delta t \to 0$, à l'évolution de $P_{n\,+}$. Au total, pour $P_{n\,+}$, le bilan de probabilité s'exprime comme :

$$P_{n\,+}(t+\Delta t) = (1-R_-\Delta t)P_{n\,+}(t) + R_+\Delta t P_{n-1\,-}(t) + \mathcal{O}(\Delta t^{k\geq 2}) \ ; \tag{21.91}$$

de même, pour $P_{n\,-}$, on obtient :

$$P_{n\,-}(t+\Delta t) = (1-R_+\Delta t)P_{n\,-}(t) + R_-\Delta t P_{n-1\,+}(t) + \mathcal{O}(\Delta t^{k\geq 2}) \ . \tag{21.92}$$

Après division par Δt, ces équations donnent dans la limite $\Delta t \to 0$ ($n = 0, 1, 2, \ldots$) :

$$\dot{P}_{n\,+}(t) = -R_-P_{n\,+} + R_+P_{n-1\,-} \ , \qquad \dot{P}_{n\,-}(t) = -R_+P_{n\,-} + R_-P_{n-1\,+} \ ; \tag{21.93}$$

en particulier, comme $P_{-1\,\pm} = 0$:

$$\dot{P}_{0\,\pm}(t) = -R_{\mp}P_{0\,\pm}(t) \iff P_{0\,\pm}(t) = C\,\mathrm{e}^{-R_{\mp}t} \ . \tag{21.94}$$

Supposons que le signal de fluorescence s'éteint à l'instant t_0, signifiant que l'atome vient d'arriver dans l'état métastable ; la probabilité qu'il n'y ait eu aucune transition entre $t = t_0$ et $t = t_0$ et que l'atome soit dans l'état + est égale à 1, d'où $P_{0\,+}(0) = 1$, donnant :

$$P_{0\,+}(t) = \mathrm{e}^{-R_-t} \ . \tag{21.95}$$

Ce dernier résultat peut d'ailleurs se retrouver immédiatement comme suit. $P_{0\,+}(t)$ est finalement la probabilité pour que le système étant parti de $|+\rangle$ ($= |\mathrm{m}\rangle$) à l'instant zéro, soit en t dans le même état *sans avoir jamais transité*, autrement dit la probabilité pour que l'atome ait été en permanence dans $|+\rangle$ entre 0 et t. La probabilité de transition par unité de temps pour le passage $|\mathrm{m}\rangle \to |\mathrm{g}\rangle$ est R_- ; on peut donc écrire de façon évidente :

$$P_{0\,+}(t+\Delta t) = (1-R_-\Delta t)P_{0\,+}(t) + \mathcal{O}(\Delta t^{k\geq 2}) \ , \tag{21.96}$$

d'où l'on déduit par passage à la limite :

$$\dot{P}_{0\,+}(t) = -R_-\,P_{0\,+}(t) \iff P_{0\,+}(t) = \mathrm{e}^{-R_-t} \ . \tag{21.97}$$

Cette expression donne donc la probabilité pour que le signal de fluorescence soit encore éteint à l'instant $t+t_0$ quand il l'était à $t = t_0$ – quel que soit l'instant t_0 puisque l'on est en régime permanent – sans s'être jamais allumé dans l'intervalle ; évidemment, plus le temps passe, plus cette probabilité est petite. Si T désigne la variable aléatoire : durée de la période noire, on a donc :

$$\mathrm{Prob}[\,T > t\,] = \mathrm{e}^{-R_-t} \ , \tag{21.98}$$

d'où la fonction de répartition[46] de T : $1 - \mathrm{e}^{-R_- t}$. On en déduit la densité de probabilité de la durée des périodes noires :

$$\lim_{\delta t \to 0} \frac{1}{\delta t} \mathrm{Prob}[\, t < T \leq t + \delta t \,] = R_- \, \mathrm{e}^{-R_- t} \ . \tag{21.99}$$

Les temps de résidence dans l'état métastable[47] sont donc théoriquement distribués suivant une loi exponentielle, qui est bien vérifiée expérimentalement [118] moyennant quelques corrections venant affiner après coup le modèle théorique de Cook et Kimble. La connaissance de la loi (21.99) permet notamment d'obtenir la durée moyenne des périodes sans signal :

$$\langle T \rangle = \int_0^{+\infty} t \, R_- \, \mathrm{e}^{-R_- t} \, \mathrm{d}t = \frac{1}{R_-} \ . \tag{21.100}$$

Des arguments analogues montrent que la durée des périodes "blanches" est distribuée selon la densité $R_+ \, \mathrm{e}^{-R_+ t}$ et que sa moyenne est $\frac{1}{R_+}$. La durée des périodes noires est donc distribuée selon une loi exponentielle donnée par (21.99), à nouveau assez bien vérifié expérimentalement [118].

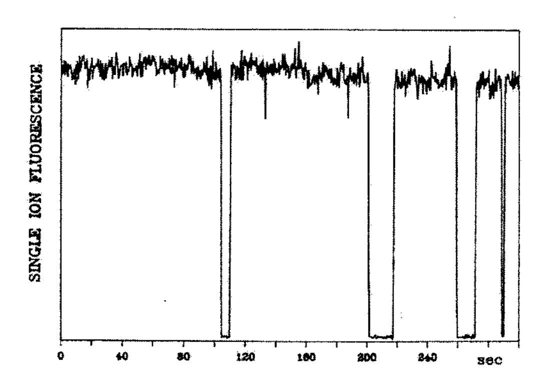


Figure 21.9: Signal de fluorescence intermittente d'*un* ion Ba^+ (figure extraite de [110]) – comparer avec fig. 21.7.

Il est également possible de calculer (et de mesurer) la fonction d'autocorrélation (centrée) de l'intensité, qui donne des renseignements plus fins sur le processus aléatoire. Elle est définie comme :

$$C(t) \stackrel{\text{déf}}{=} \langle I(t_0) I(t_0 + t) \rangle - \langle I(t_0) \rangle \langle I(t_0 + t) \rangle = \langle I(t_0) I(t_0 + t) \rangle - I_0^2 \ ; \tag{21.101}$$

[46]On rappelle que la fonction de répartition est l'outil fondamental de la théorie des probabilités. Pour une variable aléatoire continue X prenant ses valeurs dans $\mathbb{R}$, cette fonction est (le plus souvent) définie comme égale à la probabilité pour que X soit inférieure ou égale à la valeur donnée x, soit $F(x) \stackrel{\text{déf}}{=} \mathrm{Prob}[X \leq x]$; $F(x)$ est une fonction positive non-décroissante, continue à droite et telle que $F(-\infty) = 0$, $F(+\infty) = 1$. La considération de $F(x)$ permet d'inclure les variables singulières continues (*théorème de décomposition* de Lebesgue), dont la plus célèbre est celle admettant la fonction de Cantor comme loi de répartition : c'est une variable uniformément distribuée sur l'ensemble (*poussière*) de Cantor.

[47]On appelle *shelving* le phénomène de piégeage dans cet état.

comme la variable aléatoire I vaut 0 ou I_0, le produit $I(t)I(t')$ peut prendre deux valeurs, 0 ou I_0^2 – cette dernière valeur étant obtenue si le signal existe à l'instant t *et* à l'instant t'. L'espérance mathématique du produit contient donc un seul terme :

$$\langle I(t)I(t')\rangle = I_0^2\, P_{\text{on on}}(t,\, t') \ ; \qquad (21.102)$$

où $P_{\text{on on}}(t,\, t')$ est la probabilité pour que le signal soit "on" à l'instant t et à l'instant t'. Grâce à la stationnarité, il suffit de trouver $P_{\text{on on}}(t_0,\, t_0+t) \equiv P_{\text{on on}}(t)$, probabilité pour que l'on soit dans une période blanche à l'instant t_0 *et* à l'instant t_0+t, quel que soit le nombre de transitions intervenues entre t_0 et t_0+t (on peut donc définitivement poser $t_0 = 0$).

Pour trouver cette probabilité, il suffit de faire à nouveau le bilan des probabilités entre t et $t+\Delta t$. C'est R_+ qui éteint la fluorescence, donc $1-R_+\Delta t$ donne la probabilité la fluorescence reste allumée Δt plus tard. Il y a un autre terme dans le bilan, introduisant la probabilité $P_{\text{on off}}$ pour que le système soit "on" à t et "off" à 0. En effet, *via* R_- – qui allume la fluorescence – il y a un autre chemin contribuant à $P_{\text{on on}}(t)$: c'est celui où le sytème est allumé en 0, éteint à t et bascule entre t et $t+\Delta t$. Le bilan est donc :

$$P_{\text{on on}}(t+\Delta t) = (1-R_+\Delta t)\, P_{\text{on on}}(t) + R_-\Delta t\, P_{\text{on off}}(t) \ . \qquad (21.103)$$

Il faut donc aussi connaître $P_{\text{on off}}(t)$; des arguments analogues montrent que cette probabilité satisfait :

$$P_{\text{on off}}(t+\Delta t) = (1-R_-\Delta t)\, P_{\text{on off}}(t) + R_+\Delta t\, P_{\text{on on}}(t) \ . \qquad (21.104)$$

Dans la limite $\Delta t \rightarrow 0$, on obtient :

$$\dot{P}_{\text{on on}}(t) = -R_+P_{\text{on on}} + R_-P_{\text{on off}} \ , \qquad \dot{P}_{\text{on off}}(t) = -R_-P_{\text{on off}} + R_+P_{\text{on on}} \ . \quad (21.105)$$

Lorsque le processus a atteint sa phase stationnaire[48], $P_{\text{on on}}(t=0) = \mathcal{P}_{-\,\text{eq}} = \frac{R_-}{R_++R_-}$, d'après (21.89) ; par ailleurs, $P_{\text{on off}}(t=0) = 0$. La solution du système (21.105 avec ces conditions initiales est :

$$P_{\text{on on}}(t) = \frac{R_-^2}{(R_++R_-)^2} + \frac{R_+R_-}{(R_++R_-)^2}\,\mathrm{e}^{-(R_++R_-)t} \ . \qquad (21.106)$$

Le report de cette expression dans (21.101) donne :

$$C(t) = \sigma_I^2\ \mathrm{e}^{-(R_++R_-)t} \qquad (t>0) \ , \qquad (21.107)$$

avec :

$$\langle I\rangle = \frac{R_-}{R_++R_-}\,I_0 \ , \qquad \sigma_I^2 \overset{\text{déf}}{=} \langle I^2\rangle - \langle I\rangle^2 = \frac{R_+R_-}{(R_++R_-)^2}\,I_0^2 \ . \qquad (21.108)$$

[48] où, notamment, toutes les fonctions de corrélation $\langle A(t)A(t')\rangle$ ne dépendent que de la *différence* des temps.

Pour un processus aléatoire stationnaire, toutes les fonctions d'autocorrélation sont paires[49] ; l'expression (21.107) est donc vraie $\forall\, t$ à condition juste d'y remplacer t par $|t|$. La parité de $C(t)$ entraîne que le spectre $S(\omega)$ des fluctuations d'intensité est une fonction à valeurs réelles :

$$S(\omega) \equiv \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{d}t\, C(t)\, \mathrm{e}^{\mathrm{i}\omega t} = \int_{0}^{+\infty} \mathrm{d}t\, C(t)\, \cos\omega t = \frac{\sigma_I^2\,(R_+ + R_-)}{\omega^2 + (R_+ + R_-)^2} \ . \tag{21.109}$$

C'est une lorentzienne de largeur $\Delta\omega = R_+ + R_-$.

La possibilité d'observer un très petit nombre d'atomes ou ions a ainsi permis de mettre à jour les mécanismes élémentaires imaginés par Bohr. En outre, notamment pour des raisons méthodologiques, elle a ouvert une nouvelle voie théorique construite sur la notion de fonction d'onde stochastique [119]. En effet, la description usuelle des processus optiques dissipatifs consiste à écrire une équation maîtresse (équations de Bloch optiques) pour l'opérateur-densité relatif à un ensemble statistique d'atomes, bien approché en pratique par une vapeur contenant un très grand nombre d'entités élémentaires *indépendantes*. En-dehors des difficultés techniques à surmonter pour obtenir la solution d'une telle équation[50], il est bien clair que ce formalisme est très mal adapté au cas des petits amas et, à la limite, au cas de l'atome unique – on commence par calculer des propriétés d'ensemble pour essayer d'en déduire des comportements individuels, ce qui recourt implicitement à une hypothèse du genre ergodique. Un traitement plus direct est donc visiblement souhaitable.

Le choix de se cantonner à un formalisme *purement mécanique* pose le problème de la description de l'irréversibilité (dissipation). Comme déjà mentionné, celle-ci apparaît naturellement quand on couple un petit système à un grand système, et est *a contrario* exclue dans un formalisme hamiltonien standard concernant un petit nombre de degrés de liberté et en l'absence de toute trace partielle. S'en tenir à une description en termes de fonction d'onde passe donc par la réconciliation de deux concepts antagonistes, en *forçant* la mécanique à être irréversible. Une façon d'induire l'irréversibilité est de ponctuer la dynamique par des opérations de réduction de paquets d'ondes, c'est-à-dire d'effectuer par la pensée des mesures idéales, entrant dans la catégorie des *gedanken experiments*.

Ces idées vont maintenant être illustrées dans le cas [119] d'un atome à deux niveaux couplé au champ électromagnétique quantifié, l'atome étant par ailleurs piloté par une onde monochromatique de pulsation ω, traitée classiquement et représentée par $\mathcal{E}_0 \cos\omega t$. Pour simplifier, on supposera que le champ quantifié peut être soit dans l'état fondamental $|0\rangle$ (vide de photons), soit dans un état à un photon d'impulsion $\hbar\vec{k}$ et de polarisation ε, état noté $|\hbar\vec{k}\,\varepsilon\rangle$. Les deux états atomiques "nus" sont notés $|\mathrm{g}\rangle$ et $|\mathrm{e}\rangle$ et ont pour énergies E_g et $E_\mathrm{e} = E_\mathrm{g} + \hbar\omega_0$. Le champ classique fait osciller l'atome *à la*

[49]Ceci n'est vrai que pour des grandeurs *classiques*, et résulte du fait que tous les produits commutent. Dans un cadre quantique, même pour un processus stationnaire, une fonction de corrélation $\langle A(t)B(t')\rangle$ n'a plus cette propriété sauf si les opérateurs $A(t)$ et $B(t')$ commutent quels que soient les instants t et t', ou si la fonction de corrélation est symétrisée en $(1/2)\langle A(t)B(t') + B(t')A(t)\rangle$.

[50]Pour un système à N niveaux, l'opérateur densité est représenté par une matrice hermitique $N \times N$; il faut donc calculer, engranger et manipuler environ N^2 quantités réelles.

Rabi ; en définissant le zéro d'énergie à la valeur $E_{\rm g}$, le Hamiltonien effectif décrivant l'oscillation de Rabi est[51] :

$$H_{\rm Rabi} = \hbar(\omega_0 - \omega)\,|e\rangle\langle e| + \frac{\hbar\Omega}{2}(|e\rangle\langle g| + |g\rangle\langle e|) \ , \qquad (21.110)$$

où $\Omega = -d\mathcal{E}_0$, d étant le moment dipolaire électrique de l'atome (Ω est la fréquence de Rabi) ; $\delta = \omega - \omega_0$ est le désaccord entre le laser et la pulsation atomique. Quand $\delta = 0$, le Hamiltonien (21.110) induit bien un mouvement où les probabilités pour l'atome d'être dans ses deux états alternent en prenant successivement dans le temps les deux valeurs extrêmes 0 et 1 (voir problème 18.7.11).

Supposons qu'à l'instant t, le système complet (atome + champ quantifié) est dans l'état factorisé :

$$|\Psi(t)\rangle = (a_{\rm g}|g\rangle + a_{\rm e}|e\rangle) \otimes |0\rangle \equiv |\Phi(t)\rangle \otimes |0\rangle \ . \qquad (21.111)$$

Entre[52] t et $t+\delta t$, il y a possibilité d'émission spontanée de photons ; la fonction d'onde à l'instant $t+\delta t$ peut se décomposer en la somme de deux termes, chacun d'entre eux correspondant à un nombre donné, 0 ou 1, de photons du champ quantifié :

$$|\Psi(t+\delta t)\rangle = |\Psi^{(0)}(t+\delta t)\rangle + |\Psi^{(1)}(t+\delta t)\rangle \qquad (21.112)$$

avec[53] :

$$|\Psi^{(0)}(t+\delta t)\rangle = (a'_{\rm g}|g\rangle + a'_{\rm e}|e\rangle) \otimes |0\rangle \ , \quad |\Psi^{(1)}(t+\delta t)\rangle = |g\rangle \otimes \sum_{\vec{k},\,\varepsilon} \beta_{\vec{k},\,\varepsilon}\,|\vec{k},\,\varepsilon\rangle \ . \qquad (21.113)$$

Les expressions précises de ces états peuvent être obtenues de diverses façons (Wigner - Weisskopf, résolvante dans l'approximation du pôle, Hamiltonien effectif non-hermitique (voir p. 1167), etc.). Le carré de la norme de $|\Psi^{(1)}\rangle$ est la probabilité δp d'émission d'un photon entre t et $t+\delta t$, soit $\delta p = \Gamma_{\rm e}|a_{\rm e}|^2\delta t$; l'intervalle de temps δt est supposé court au sens où $\delta p \ll 1$. Par la conservation de la probabilité totale, il en résulte que $|\langle\Psi^{(0)}|\Psi^{(0)}\rangle|^2 = 1 - \delta p$. Les amplitudes $a'_{\rm g}$ et $a'_{\rm e}$ s'obtiennent en faisant évoluer l'atome avec $H \equiv H_{\rm Rabi} - {\rm i}\hbar(\Gamma_{\rm e}/2)\,|{\rm e}\rangle\langle{\rm e}|$; enfin, $\beta_{\vec{k},\,\varepsilon}$ est l'amplitude de probabilité d'émission d'un photon d'impulsion $\hbar\vec{k}$ et de polarisation ε. Au total, l'évolution du système total est unitaire (la norme de $|\Psi(t+\delta t)\rangle$ est la même, par construction, que celle de $|\Psi(t)\rangle$).

C'est maintenant que s'introduit l'irréversibilité induite par l'"observation". On suppose que tous les photons émis spontanément sont détectés avec un compteur parfait, qui mesure le nombre de photons dans le champ à l'instant $t+\delta t$; tout photon ainsi mesuré est absorbé par le compteur et disparaît donc du champ électromagnétique. Conformément au postulat de réduction du paquet d'ondes, si le résultat est 0, on doit

[51]On s'en tient à l'approximation de l'onde tournante (*Rotating Wave Approximation*, acronyme RWA), où les termes antirésonnants en $\mathrm{e}^{\pm 2{\rm i}\omega t}$ sont négligés.

[52]δt est supposé très petit devant $\Gamma_{\rm e}^{-1}$, durée de vie de l'état excité, δ^{-1} et Ω^{-1}.

[53]Dans le secteur à zéro photon, l'atome peut être soit en $|g\rangle$, soit en $|e\rangle$ (transitions virtuelles sans évacuation du photon). Dans le secteur à 1 photon, l'atome est dans l'état fondamental, puisqu'un photon a été émis vers l'extérieur. Par ailleurs, la désexcitation suivie d'une réexcitation pendant δt est un processus à deux étapes qui est d'ordre δt^2.

projeter l'état juste avant la mesure, $|\Psi(t+\delta t)\rangle$, sur $|\Psi^{(0)}(t+\delta t)\rangle$, si au contraire on trouve 1, on projette sur $|\Psi^{(1)}(t+\delta t)\rangle$; dans tous les cas, il faut renormaliser l'état ainsi obtenu puisqu'une deuxième mesure effectuée immédiatement après la première doit donner le même résultat avec probabilité 1.

Dans la procédure numérique, le caractère aléatoire de l'émission spontanée pendant le petit intervalle δt est très facile à singer : on se fixe un seuil $\eta \ll 1$ et on tire une suite de nombres au hasard r_i répartis uniformément, entre 0 et 1 par exemple. $0 < r_i \leq \eta$ correspond à l'émission d'un photon, alors que $\eta < r_i < 1$ correspond à la non-émission. Au total, *après* cette mesure, l'alternative est la suivante :

- soit $r_i < \eta$, auquel cas :

$$|\Psi(t+\delta t)\rangle = |g\rangle \otimes \sum_{\vec{k},\,\varepsilon} \beta_{\vec{k},\,\varepsilon} |\vec{k},\,\varepsilon\rangle \ , \tag{21.114}$$

- soit $r_i > \eta$, auquel cas[54] :

$$|\Psi(t+\delta t)\rangle = \mu\,(a'_{\mathrm{g}}|g\rangle + a'_{\mathrm{e}}|e\rangle) \otimes |0\rangle = \mu\,(\mathbf{1} - \mathrm{i}\hbar H\delta t)|\Phi(t)\rangle \otimes |0\rangle \ , \tag{21.115}$$

avec $\mu = (1-\delta p)^{-\frac{1}{2}}$. En tout cas, à l'issue de chaque mesure, on en revient à un état factorisé de la forme (21.111) où le champ est dans l'état vide, et le cycle peut être recommencé. Dans ce schéma, l'évolution de la fonction d'onde est une marche au hasard forcée dans l'espace des états atomiques, d'où le nom de *fonction d'onde stochastique*.

La question se pose de l'équivalence entre l'approche stochastique et la description conventionnelle à l'aide des équations de Bloch optiques. Dalibard et al. [119] ont également montré que la moyenne d'ensemble de l'opérateur densité construit sur la fonction d'onde stochastique satisfait une équation d'évolution identique à celle de Bloch. D'ailleurs, les simulations numériques [119] montrent qu'avec une centaine de tirages (d'histoires), les résultats obtenus par l'approche stochastique sont quasiment indiscernables de ceux fournis par les équations de Bloch. Même sans rentrer dans une comparaison détaillée des deux procédures, il est clair que la description stochastique est bien moins coûteuse. La situation est analogue à celle que l'on rencontre dans la théorie conventionnelle du mouvement Brownien ; alors l'alternative est la suivante : ou bien on résout une équation aux dérivées partielles (équation de Fokker - Planck) pour la fonction de densité[55] – ce qui permet d'avoir accès à toutes les moyennes en tant qu'espérances mathématiques –, ou bien on simule numériquement la trajectoire brownienne, ce qui est en principe très facile, et on moyenne sur un (très grand) nombre de trajectoires (histoires). Dans tous les cas, la dynamique moyennée est foncièrement irréversible : qu'elle se produise dans l'espace physique ou dans un espace plus abstrait, une marche au hasard présente toujours un comportement où la symétrie de renversement du temps est brisée[56].

[54]Après mesure, le photon émis a été extrait du système par le compteur et le champ est bien à nouveau dans l'état vide.

[55]définie dans l'espace des phases, en général.

[56]Essentiellement, la brisure de symétrie provient du fait que l'on s'intéresse à un petit nombre de

21.3 Cryptographie quantique

La cryptographie est l'art – vieux comme le monde[57] – de communiquer au moyen de conventions assurant que le message transmis est à l'abri des indiscrétions. L'idée la plus naturelle est d'utiliser un *codage* des messages confidentiels : toute lettre de l'alphabet est remplacée biunivoquement par un autre signe (nombre – d'où l'expression message chiffré –, image, symbole exotique, etc.). La formule permettant de constituer le lexique donnant la correspondance biunivoque s'appelle la clé, et c'est elle qui doit rester absolument confidentielle ; portant précisément le nom de *cryptographie à clé secrète*[58], ce procédé de chiffrage est visiblement le plus simple qui soit. Il existe de multiples recettes plus élaborées ; par exemple, on code une lettre en langage binaire, lui associant une suite $\{x_i\}$ de 0 ou de 1, laquelle n'a rien de confidentiel (tout le monde peut prendre la même) ; puis on tire au hasard une autre suite $\{y_i\}$ de 0 ou de 1, qui constitue la clé à proprement parler et qui, elle, est confidentielle. Le chiffrage de la lettre choisie est constitué par la suite des sommes $\{x_i + y_i \text{ modulo } 2\}$. Quoi qu'il en soit, l'élément crucial est la *confidentialité* de la clé permettant de coder et de décoder le message échangé, étant entendu que celui-ci peut à la limite être rendu public. Tout le problème est donc d'assurer la sécurité du message initial définissant la clé.

Toute transmission classique d'information peut être interceptée passivement, c'est-à-dire sans modification de la teneur du message ; bien sûr, il peut y avoir d'énormes difficultés techniques à surmonter pour y parvenir, mais c'est en principe possible : l'écoute d'une ligne téléphonique perturbe en un sens la ligne – donc ouvre la voie à la révélation de l'écoute – mais l'espion peut tenter de réduire à l'extrême cet effet révélateur pour rendre son intervention quasiment indétectable. Au contraire, parce qu'une mesure quantique perturbe en général[59] le système objet de la mesure, l'observation indésirable

degrés de liberté, relatifs à un sous-système du système considéré. Dans le cas de la particule brownienne, on regarde sa position, sa vitesse, etc., étant entendu que le "bain" – qui provoque la marche au hasard – est complètement évacué d'emblée et n'est présent que par son effet erratique sur la marche (par une équation de Langevin). Un autre exemple est fourni par l'amortissement de Landau dans un plasma : l'équation de départ (de Vlasov, où le terme de collisions est absent, au contraire de l'équation de Boltzmann) est réversible et produit une densité $f(\vec{r}, \vec{p}, t)$ permettant de calculer les valeurs moyennes. Si l'on s'intéresse seulement à la densité de particules, il convient d'intégrer sur les impulsions, ce qui fait effacer une grande quantité d'information ; l'irréversibilité qui en découle traduit une évidence : une fois effectuée la sommation sur les vitesses, il n'est plus possible de reconstruire le mouvement, il y a une infinité de conditions initiales susceptibles de donner la même densité en position. Sommer, intégrer ou laisser dans l'ombre certains degrés de liberté en les confiant au hasard produit ici la flèche du temps.

[57]On raconte que Jules César y a recouru : était-ce pour correspondre avec Cléopâtre ?

[58]Il existe aussi des procédés dits *à clé publique*, l'idée étant que certaines opérations mathématiques s'effectuent trivialement dans un sens et exigent un travail colossal dans l'autre (pour cette raison, on parle de *fonctions à sens unique*) : par exemple, il est facile de multiplier deux nombres premiers, alors que la factorisation d'un (grand) nombre N en ses premiers entiers exige un temps de calcul croissant en gros exponentiellement avec N. Typiquement, factoriser un nombre de mille chiffres en ses facteurs premiers est, avec les ordinateurs actuels, totalement inconcevable... du moins si l'on s'en tient aux algorithmes connus publiquement (nul ne devrait affirmer qu'il n'existe pas quelque part un algorithme tenu secret permettant de *casser* un code avec les moyens de calcul actuels...).

C'est d'ailleurs pour ce type de problème que l'ordinateur quantique repointe le bout du nez, compte tenu de ses capacités *potentielles* phénoménales...

[59]La mesure ne provoque pas *toujours* une *perturbation incontrôlable*, comme on le voit parfois écrit. Si l'état du système est, avant la mesure, propre de l'observable mesurée, il l'est encore après : en pareil

peut être en principe facilement détectée. La Mécanique quantique permet ainsi de construire des procédures permettant de savoir *avec certitude* si, lors d'une transmission, on est écouté ou non. Il s'agit ici de se borner à donner quelques idées de base, étant entendu qu'il existe une vaste littérature sur le sujet, où divers scénarios sont envisagés et décrits[60]. À la suite de Rivest [122], l'habitude a été prise d'appeler Alice et Bob les deux personnes communiquant entre elles et souhaitant se mettre à l'abri des malversations d'un "espion", souvent dénommé Eve (vient de *to eavesdrop* = *écouter aux portes* (!)).

Dans la suite, on considère un premier exemple, formel, posant les idées essentielles montrant comment la Mécanique quantique permet de concevoir des protocoles révélant *à coup sûr* la présence d'une écoute indésirable. Deux exemples utilisant soit un atome à trois niveaux, soit des spins 1/2 sont ensuite exposés en détail. Dans un dernier temps, on examine le point essentiel, à savoir comment transmettre la clé en toute confidentialité, suivant les propositions ingénieuses de Bennett et Brassard (BB84) [123], et d'Ekert [124].

21.3.1 Principes de la détection infaillible d'une écoute indésirable

L'idée de base est simple et met en jeu des séquences de mesures successives d'observables qui ne commutent pas (voir Tome I, problème 12.3.5). De façon un peu formelle, les choses se présentent comme suit. Soit A et B deux observables incompatibles :

$$[A,\, B] \neq 0 \ , \tag{21.116}$$

dont l'une, A pour fixer les idées, est une constante du mouvement :

$$[A,\, H] = 0 \quad , \tag{21.117}$$

alors que B n'en est pas une ($[B,\, H] \neq 0$). À un certain instant, $t = 0$, on mesure A, et on trouve a_m, l'une des valeurs propres de A ; on sait d'avance que la probabilité d'obtenir cette valeur est égale à $|\langle a_m|\Psi(t=0)\rangle|^2$. Juste après cette mesure, l'état du système est $|\Psi(t = 0+0)\rangle = |a_m\rangle$. Si aucune autre mesure n'est effectuée[61] jusqu'à l'instant t, une nouvelle mesure de A à cet instant donne le résultat a_m, cette fois avec certitude. En effet, l'état à l'instant t est $|\Psi(t)\rangle = U(t)|\Psi(t = 0+0)\rangle = U(t)|a_m\rangle$, mais comme A et H commutent, l'état $|a_m\rangle$ produit par la réduction du paquet d'onde consécutive à la première mesure de A, est également[62] propre de H ; il en résulte que $|\Psi(t)\rangle = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_m t}\,|a_m\rangle$ et que la deuxième mesure de A produit la même valeur a_m avec probabilité 1. Si les deux mesures sont effectuées par deux observateurs distincts (ceux que l'usage appelle Alice et Bob), ces derniers doivent donc constater que leurs résultats sont toujours identiques, quand la même séquence est répétée un grand nombre de fois,

cas, la mesure est intrinsèquement *non-destructive*, laissant le système exactement dans le même état qu'avant la mesure.

[60] Pour une bibliographie assez récente, voir [120] ; pour un point sur la question, voir [121].

[61] Dans ces conditions, le système évolue librement par l'équation de Schrödinger.

[62] Pour simplifier la discussion, on suppose pour l'instant qu'il n'existe aucune dégénérescence.

à la limite infini. Techniquement, ceci se traduit par l'égalité donnant la probabilité conditionnelle pour que Bob trouve la valeur $a_{m'}$ sachant qu'Alice a trouvé a_m :

$$\boxed{P(a_{m'}|a_m) = \delta_{mm'}} \tag{21.118}$$

Imaginons maintenant au contraire que, entre la première mesure de A faite par Alice et la seconde effectuée par Bob, une tierce personne mesure à t_1 ($0 < t_1 < t$) l'observable B, qui n'est pas une constante du mouvement. Cet *espion* (Eve) trouve l'une des valeurs propres de B, b_n, et l'état issu de cette mesure est[63] l'état projeté selon le postulat de réduction du paquet d'ondes $|\Psi(t = t_1 + 0)\rangle = |b_n\rangle$. Il s'agit d'un état non-stationnaire, puisque H et B ne commutent pas ; à l'instant ultérieur $t > t_1$, cet état est devenu :

$$|\Psi(t)\rangle = \sum_{m'} c_{m'}\, \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_{m'}(t-t_1)}\, |a_{m'}\rangle \ , \qquad c_{m'} = \langle a_{m'}|b_n\rangle, \qquad 0 < t_1 < t \ . \tag{21.119}$$

Dans ces conditions, la mesure de A par Bob à l'instant t postérieur à t_1 peut produire *l'une quelconque* des valeurs propres $a_{m'}$ (la probabilité de trouver $a_{m'}$ est $P_{m'} = |c_{m'}|^2$) ; techniquement, ceci se traduit maintenant par[64] :

$$\boxed{P(a_{m'}|a_m) = \sum_n P^{**}(a_{m'}|b_n)P^*(b_n|a_m)} \tag{21.120}$$

où $P^*(b_n|a_m)$ est la probabilité conditionnelle que l'espion trouve b_n quand Alice a trouvé a_m, $P^{**}(a_{m'}|b_n)$ celle que Bob trouve $a_{m'}$ quand l'espion a trouvé b_n.

En définitive, les résultats obtenus par Alice et Bob effectuant les deux mesures de A ne sont plus nécessairement identiques. En fait, lorsque la séquence est répétée un certain nombre de fois, il suffit que l'*un* de leurs résultats soit différent pour qu'ils puissent affirmer avec certitude qu'ils sont victimes d'une écoute et que la sécurité de leur transmission n'est pas assurée.

21.3.2 Exemples

Exemple 1 : trois observables

Précisons les idées précédentes dans un cas concret, où d'ailleurs le Hamiltonien présente de la dégénérescence. Soit un atome à trois niveaux $\{|p\rangle\}$ (un fondamental, $p = 0$, et deux niveaux excités dégénérés, $p = 1, 2$) :

$$H = \hbar\omega(-|0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|) \ . \tag{21.121}$$

[63] On continue à supposer qu'il n'existe pas de dégénérescence. L'exemple 1 ci-dessous présente un cas dégénéré.

[64] C'est la relation de chaîne, dite *de Bachelier - Chapman - Kolmogorov*. La somme représente l'addition des probabilités d'événements mutuellement exclusifs. C'est cette décomposition qui intervient dans l'expérience d'Young quand on sait par quelle fente est passé l'électron (voir Tome I, section 10.1).

A est une observable donnée par :

$$A = a\,(|0\rangle\langle 0| + \mathrm{i}|1\rangle\langle 2| - \mathrm{i}|2\rangle\langle 1|) \qquad (a \in \mathbb{R}) \ . \tag{21.122}$$

A commute visiblement avec H, c'est donc une constante du mouvement. Ses valeurs propres sont $+a$ (dégénérescence 2) et $-a$ (non-dégénérée). Les vecteurs propres de A (qui sont aussi propres de H sont les $|\psi_{\alpha\beta}\rangle$ $(\alpha,\, \beta = \pm)$:

$$|+a,\, 1\rangle = |0\rangle \equiv |\psi_{-+}\rangle \ , \qquad |+a,\, 2\rangle = \frac{1}{\sqrt{2}}(|1\rangle - \mathrm{i}|2\rangle) \equiv |\psi_{++}\rangle \ . \tag{21.123}$$

$$|-a\rangle = \frac{1}{\sqrt{2}}(-\mathrm{i}|1\rangle + |2\rangle) \equiv |\psi_{+-}\rangle \ . \tag{21.124}$$

Le premier indice de $|\psi_{\alpha\beta}\rangle$ renvoie à la valeur propre $\pm\hbar\omega$ de H, le second au signe de la valeur propre $\pm a$ de A. Le couple $(H,\, A)$ forme visiblement un ECOC puisqu'il n'existe qu'un seul vecteur propre associé à un couple $(\alpha,\, \beta)$ donné.

L'observable B est prise sous la forme :

$$B = b\,(|0\rangle\langle 1| + |1\rangle\langle 0| + |2\rangle\langle 2|) \qquad (b \in \mathbb{R}) \ ; \tag{21.125}$$

clairement :

$$[A,\, B] \neq 0 \ , \qquad [H,\, B] \neq 0 \ . \tag{21.126}$$

Les valeurs propres de B sont $+b$ (dégénérescence 2) et $-b$ (non-dégénérée) ; les vecteurs propres sont :

$$|+b,\, 1\rangle = |2\rangle \ , \qquad |+b,\, 2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \ , \qquad |-b\rangle = \frac{1}{\sqrt{2}}(-|0\rangle + |1\rangle) \ . \tag{21.127}$$

Maintenant, Alice et Bob effectuent des mesures de A aux instants 0 et t. Compte tenu de la dégénérescence en énergie, l'opération de mesure peut porter sur un ensemble complet d'observables qui commutent, ce que constituent A et H ; à l'issue d'une telle opération, l'état du système est parfaitement défini. Si aucune mesure de B n'est faite entre-temps, les résultats des deux observateurs coïncident dans tous les cas.

En revanche, soit une mesure de B effectuée à l'instant t_1 compris entre 0 et t. Pour fixer les idées, on suppose que le premier observateur a trouvé $+\hbar\omega$ et $+a$; l'état de départ est donc $|\Psi(0+)\rangle = |\psi_{++}\rangle$. Compte tenu des décompositions ci-dessus, l'état à l'instant t_1 est :

$$|\Psi(t_1)\rangle = \mathrm{e}^{-\mathrm{i}\omega t_1}\,\frac{1}{\sqrt{2}}\Big[\frac{1}{\sqrt{2}}(|+b,\, 2\rangle + |-b\rangle) - \mathrm{i}|+b,\, 1\rangle\Big] \ . \tag{21.128}$$

Il en résulte que la mesure de B à t_1 donne $+b$ avec la probabilité $\frac{1}{4} + \frac{2}{4} = \frac{3}{4}$, $-b$ avec la probabilité $\frac{1}{4}$. Dans le premier cas, l'état à $t_1 + 0$ est :

$$|\Psi(t_1 + 0)\rangle = \frac{\sqrt{2}}{\sqrt{3}}\Big[\frac{1}{\sqrt{2}}|+b,\, 2\rangle - \mathrm{i}|+b,\, 1\rangle\Big] \ . \tag{21.129}$$

En exprimant les états propres de B sur les $|\psi_{\alpha\beta}\rangle$, on voit aisément que cet état devient à l'instant t :

$$|\Psi(t)\rangle = \frac{\sqrt{2}}{2\sqrt{3}}\left[\mathrm{e}^{\mathrm{i}\omega(t-t_1)}|\psi_{-+}\rangle + \frac{3}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\omega(t-t_1)}|\psi_{++}\rangle - \frac{\mathrm{i}}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\omega(t-t_1)}|\psi_{+-}\rangle\right] \ . \qquad (21.130)$$

Pour cette *histoire*, le deuxième observateur va trouver les couples de valeurs $(-\hbar\omega,\ +a)$, $(+\hbar\omega,\ +a)$ et $(+\hbar\omega,\ -a)$ avec les probabilités respectives $\frac{1}{6}$, $\frac{3}{4}$ et $\frac{1}{12}$, au lieu de trouver à coup sûr le même résultat que le premier observateur.

Si la mesure intermédiaire donne $-b$, on a successivement :

$$|\Psi(t_1+0)\rangle = |-b\rangle \ , \qquad (21.131)$$

qui devient :

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}}\mathrm{e}^{+\mathrm{i}\omega(t-t_1)}|\psi_{-+}\rangle + \frac{1}{2}\mathrm{e}^{-\mathrm{i}\omega(t-t_1)}|\psi_{++}\rangle + \frac{\mathrm{i}}{2}\mathrm{e}^{-\mathrm{i}\omega(t-t_1)}|\psi_{+-}\rangle \ . \qquad (21.132)$$

Le deuxième observateur va trouver les valeurs $(-\hbar\omega,\ +a)$, $(+\hbar\omega,\ +a)$ et $(+\hbar\omega,\ -a)$ avec les probabilités respectives $\frac{1}{2}$, $\frac{1}{4}$ et $\frac{1}{4}$.

Ainsi, au coup par coup, les deux observateurs (Alice et Bob) peuvent trouver des résultats différents. La probabilité qu'ils trouvent des résulats identiques est la probabilité conditionnelle :

$$P_{\text{résultats identiques}} \equiv P(++\,;\, t|++;\, 0) \ , \qquad (21.133)$$

qui peut s'exprimer suivant la relation de chaîne habituelle :

$$P(++\,;\, t|++;\, 0) = \sum_{\varepsilon=\pm1} P^{**}(++\,;\, t|\varepsilon b;\, t_1)\, P^{*}(\varepsilon b\,;\, t_1|++\,;\, 0) \ . \qquad (21.134)$$

Cette dernière expression est la somme sur toutes les "trajectoires" possibles, une mesure intermédiaire ayant été faite[65]. On a donc ici :

$$P(++\,;\, t|++\,;\, 0) = \frac{3}{4}\frac{3}{4} + \frac{1}{4}\frac{1}{4} = \frac{5}{8} \qquad (21.135)$$

Le choix d'un autre résultat pour la première mesure du couple $(H,\ A)$ conduit au même type de conclusion : en raison de la mesure effectuée par un tiers, la probabilité que les résultats soient identiques est loin d'être égale à 1 ; le calcul montre qu'elle vaut $\frac{1}{2}$ si la première mesure a donné $(-\hbar\omega,\ +a)$ et $\frac{5}{8}$ si on a d'abord trouvé $(+\hbar\omega,\ -a)$. Dans tous les cas, au bout de quelques expériences, Alice et Bob peuvent avec une quasi-certitude affirmer qu'ils sont écoutés.

[65]Lorsqu'aucune mesure intermédiaire n'est faite, la probabilité est le module au carré d'une *somme d'amplitudes*. Ici, compte tenu de la réduction liée à la mesure de B à t_1, la probabilité est une *somme de probabilités*, chacune de celles-ci étant le module carré d'une amplitude.

♦ *Remarque*

Toutes les probabilités exhibées sont indépendantes du temps, puisqu'il s'agit de séquences de mesures du type AA ou ABA. En effet :

- quand il s'agit de mesures de A, toutes les réductions de paquet d'onde produisent des états stationnaires ; aucune dépendance en temps ne peut apparaître.
- en ce qui concerne la mesure de B : l'état sur lequel on mesure est propre de A et H ; comme il s'agit d'un état stationnaire, les probabilités de trouver les valeurs $\pm b$ sont indépendantes du temps t_1 auquel est faite la mesure de B. Enfin, l'état issu de la mesure de B à t_1 est $|b,\, i\,=\,1,\, 2\rangle$ ou $|-b\rangle$. Lors de la mesure de A et H à l'instant $t > t_1$, les probabilités sont de la forme $|\langle\psi_{\alpha\beta}|U(t-t_1)|b,\, i\,=\,1,\, 2\rangle|^2$ ou $|\langle\psi_{\alpha\beta}|U(t-t_1)|-b\rangle|^2$ (voir (21.123) et (21.124)) ; U agissant sur le *bra* produit une simple phase temporelle qui disparaît du module carré.

À l'inverse, une séquence de mesures du genre ABB effectuées aux instants 0, t_1 et t, donnerait des probabilités dépendant de l'intervalle de temps $t - t_1$ pour la deuxième mesure de B (voir Tome I, problème 12.3.5) ♦

Exemple 2 : Alice, Bob et des spins $\frac{1}{2}$

Une autre procédure utilise l'envoi de spins $\frac{1}{2}$ par Alice à Bob[66]. Relativement à une direction de mesure donnée, le spin $\frac{1}{2}$ est une variable binaire qui se prête donc bien au codage : on peut mettre en correspondance biunivoque l'écriture binaire d'un nombre avec une configuration de spins $++-+-+-\ldots+$. Ici, pour un Hamiltonien à symétrie sphérique, toutes les composantes S_u des spins sont des constantes du mouvement ; les observables incompatibles en jeu sont bien évidemment les différentes composantes cartésiennes de $\vec{S}$, satisfaisant les relations caractéristiques d'un moment cinétique :

$$\vec{S}\times\vec{S} = \mathrm{i}\hbar\vec{S} \quad\Longleftrightarrow\quad [S_x,\, S_y] = \mathrm{i}\hbar S_z\ ,\ \ldots \tag{21.136}$$

En vertu de la relation $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$, on peut raisonner avec les trois composantes de $\vec{\sigma}$. Par ailleurs, puisque les composantes du spin sont des constantes du mouvement, seule une mesure intermédiaire d'une composante incompatible est susceptible de modifier, lors d'une mesure ultérieure, la valeur trouvée au départ.

Alice choisit de mesurer la composante de $\vec{\sigma}$ suivant Oz, trouve donc à chaque fois ± 1 et envoie le spin mesuré à Bob. Celui-ci mesure la composante de $\vec{\sigma}$ le long d'un

[66] D'autres versions sont possibles ; par exemple, on peut disposer d'une source fabriquant des paires de spins d'état bien défini, l'un des spins étant envoyé vers Alice, l'autre vers Bob ; un espion mesurant le spin allant vers Bob va induire des différences entre les résultats des mesures des deux observateurs – voir plus loin.

Par ailleurs, au lieu de spins, on peut utiliser des photons polarisés, et les faire passer à travers des polariseurs (filtres) ; Bob et Alice trouvent ou non les mêmes résultats selon que leurs filtres respectifs sont ou non orientés de la même façon.

axe OZ faisant l'angle θ avec l'axe Oz d'Alice et situé dans le plan xOz (Alice et Bob sont d'avance convenus d'une même orientation de leurs repères locaux). L'observable que mesure Bob est donc :

$$\sigma_Z = \cos\theta\,\sigma_z + \sin\theta\,\sigma_x \ , \qquad [\sigma_z,\,\sigma_Z] = 2\mathrm{i}\,\sin\theta\,\sigma_y \tag{21.137}$$

Bien évidemment, Alice et Bob ne vont pas trouver les mêmes résultats, puisqu'ils mesurent deux observables qui ne commutent pas – sauf si $\theta = 0,\,\pi$. Les états propres de σ_Z, qui a aussi pour valeurs propres ± 1, sont :

$$|+\rangle_Z = \cos\frac{\theta}{2}\,|+\rangle_z + \sin\frac{\theta}{2}\,|-\rangle_z \ , \qquad |-\rangle_Z = -\sin\frac{\theta}{2}\,|+\rangle_z + \cos\frac{\theta}{2}\,|-\rangle_z \ , \tag{21.138}$$

comme on le vérifie facilement. Cela étant donné, on voit que si Alice a trouvé +1 – et envoie donc à Bob un spin dans l'état $|+\rangle_z$ –, celui-ci peut trouver pour σ_Z les valeurs $+1$ ou -1, avec les probabilités respectives $\cos^2\frac{\theta}{2}$ et $\sin^2\frac{\theta}{2}$. Bien sûr, si $\theta = 0$, Alice et Bob trouvent *toujours* les mêmes résultats. Si Bob mesure suivant un axe OZ tourné de $\frac{\pi}{2}$ (appelons-le Ox), alors, selon (21.138) :

$$|+\rangle_{Z=x} = \frac{1}{\sqrt{2}}(|+\rangle_z + |-\rangle_z) \ , \qquad |-\rangle_{Z=x} = \frac{1}{\sqrt{2}}(-|+\rangle_z + |-\rangle_z) \ ; \tag{21.139}$$

dans ces conditions, si Alice a trouvé $+1$, Bob a une chance sur deux de trouver $+1$, etc. : Bob ne trouve comme Alice *que* dans 50% des cas. Enfin, si Bob choisit $\theta = \pi$, lui et Alice mesurent bien des observables qui commutent ($[\sigma_z,\,-\sigma_z] = 0$!), mais leurs résultats sont *opposés* (*anticorrélés*) : si Alice a trouvé $+1$, Bob trouve -1, et inversement.

Maintenant, supposons qu'un espion, situé sur le trajet du spin, procède à des mesures suivant un axe $OEve$ faisant l'angle ϕ avec Oz ; l'observable que mesure l'espion est donc :

$$\sigma_{Eve} = \cos\phi\,\sigma_z + \sin\phi\,\sigma_x \ , \tag{21.140}$$

observable dont les vecteurs propres sont :

$$|+\rangle_{Eve} = \cos\frac{\phi}{2}\,|+\rangle_z + \sin\frac{\phi}{2}\,|-\rangle_z \ , \qquad |-\rangle_{Eve} = -\sin\frac{\phi}{2}\,|+\rangle_z + \cos\frac{\phi}{2}\,|-\rangle_z \ . \tag{21.141}$$

Pour fixer les idées, raisonnons après une mesure par Alice ayant donné $+1$, de sorte que l'espion effectue une mesure le long de l'axe $OEve$ sur l'état $|+\rangle_z$; d'après (21.141), il trouve la valeur $+1$ avec la probabilité $P^*(+1|+1) = \cos^2\frac{\phi}{2}$ et la valeur -1 avec la probabilité $P^*(-1|+1) = \sin^2\frac{\phi}{2}$. Après sa mesure, Eve fait suivre le spin à Bob, qui procède à la mesure de σ_Z, son axe OZ faisant l'angle $\theta - \phi$ avec celui de l'espion. Si Eve a trouvé $+1$, Bob trouve $+1$ ou -1 avec les probabilités respectives :

$$P^{**}(+1|+1) = \cos^2\frac{\theta-\phi}{2} \ , \qquad P^{**}(-1|+1) = \sin^2\frac{\theta-\phi}{2} \ ; \tag{21.142}$$

au contraire, si l'espion a trouvé -1, Bob trouve -1 ou $+1$ avec les probabilités :

$$P^{**}(-1|-1) = \cos^2\frac{\theta-\phi}{2} \ , \qquad P^{**}(+1|-1) = \sin^2\frac{\theta-\phi}{2} \ . \tag{21.143}$$

En définitive, quand Eve a fait une mesure intermédiaire, la probabilité pour qu'Alice et Bob aient le même résultat +1 est, toujours selon la relation de chaîne (21.120) :

$$P_{\text{résultats identiques}} \equiv P_\phi(+,\,\theta|+,\,0) = \cos^2\frac{\phi}{2}\,\cos^2\frac{\theta-\phi}{2} + \sin^2\frac{\phi}{2}\,\sin^2\frac{\theta-\phi}{2}\ . \quad (21.144)$$

En particulier, si Alice et Bob sont convenus de mesurer le long du même axe Oz ($\theta = 0$), la probabilité de coïncidence de leurs résultats est :

$$P_{\text{résultats identiques}} \equiv P_\phi(+,\,0|+,\,0) = \cos^4\frac{\phi}{2} + \sin^4\frac{\phi}{2} = \frac{1}{2}\,(1+\cos^2\phi)\ . \quad (21.145)$$

Cette probabilité n'est égale à 1 que si l'espion utilise le même axe qu'Alice et Bob, à la direction près : si $\phi = 0$, tout le monde mesure la même observable et tous les résultats coïncident ; si $\phi = \pi$, les résultats d'Alice et Eve sont anticorrélés, ainsi que ceux d'Eve et de Bob. Au total, ceux d'Alice et Bob sont toujours en coïncidence parfaite.

L'espion, ne connaissant pas la direction commune d'observation d'Alice et Bob, peut décider de tirer au hasard la sienne ; en prenant l'angle ϕ uniformément distribué entre 0 et 2π, la probabilité de coïncidence moyennée est :

$$\langle P_{\text{résultats identiques}}\rangle = \int_0^{2\pi}\frac{d\phi}{2\pi}\,\frac{1}{2}\,(1+\cos^2\phi) = \frac{3}{4}\ . \quad (21.146)$$

Cette modification ne change pas grand chose au fait que sur un grand nombre d'envois, l'espion va être démasqué à coup presque sûr.

Au total, pour faire parvenir un message de n *bits*, Alice doit envoyer Nn spins, N étant pris très grand de sorte que les fréquences statistiques puissent être confondues avec les lois limites de probabilités. Bob peut retourner à Alice tous ses spins, par une voie dont la confidentialité a été établie antérieurement ; le produit Nn étant forcément très grand – et la probabilité (21.145) étant évidemment plus petite que 1 (sauf si l'espion connaît l'angle choisi par Alice et Bob) –, les deux observateurs vont très vite s'apercevoir s'ils sont écoutés ou non. Cette possibilité offerte par la Mécanique quantique repose crucialement sur l'existence d'observables *incompatibles*, et constitue réellement la nouveauté quantique : si une seule observable était disponible (σ_z par exemple) – ou s'il n'existait que des observables compatibles –, Alice et Bob trouverait toujours le même résultat, qu'il y ait un espion ou pas.

21.3.3 Communication de la clé secrète entre Alice et Bob

Les exemples traités ci-dessus démontrent la possibilité infaillible de savoir si la ligne est écoutée ou non mais, dans une procédure de cryptage à clé secrète, n'en définissent pas pour autant un mode de transmission sûr de la clé : si être en mesure d'affirmer avec certitude que la ligne n'est pas écoutée est déjà un immense progrès, encore faut-il trouver les moyens d'échanger la clé elle-même en toute confidentialité. Ekert ([124], [125]) a proposé le scénario qui suit, exposé d'abord dans sa version simple, puis dans une version raffinée utilisant des paires de spin couplées *à la* EPR.

Dans le protocole la plus simple, Alice envoie toujours à Bob des spins polarisés. L'état de polarisation de chaque spin résulte encore d'une mesure faite par Alice et ayant produit *un* certain résultat $\pm\frac{\hbar}{2}$, noté simplement $\pm$ dans la suite. Celle-ci tire au hasard (ou selon son humeur) sa direction de mesure ; pour simplifier, on suppose que cette direction est soit Ox (notée $\leftrightarrow$), soit Oz (notée $\updownarrow$). Donc, pour chaque spin, Alice

1. tire au hasard sa direction de mesure, Ox ou Oz,
2. note son résultat, $+$ ou $-$,
3. envoie le spin à Bob (sans commentaire !).

Une séquence possible de 15 mesures par Alice est la suivante, où la deuxième ligne donne la direction de mesure choisie par Alice, la troisième étant le résultat de la mesure correspondante :

$$\begin{array}{ccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ \updownarrow & \updownarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \updownarrow & \updownarrow & \leftrightarrow \\ + & - & - & - & + & + & - & + & - & - & + & + & - & + & + \end{array} \; . \tag{21.147}$$

Après sa mesure, Alice envoie le spin à Bob. Pour chaque spin reçu, celui-ci fait une mesure en tirant lui aussi au hasard sa direction d'observation (il se trompe donc dans un cas sur deux) :

- quand il choisit le même axe qu'Alice, il trouve le même résultat qu'elle :

$$P(\text{Bob, O}z,\, +|\text{Alice, O}z,\, +) = 1 \;,\; P(\text{Bob, O}z,\, -|\text{Alice, O}z,\, +) = 0 \;, \tag{21.148}$$

$$P(\text{Bob, O}z,\, +|\text{Alice, O}z,\, -) = 0 \;,\; P(\text{Bob , O}z,\, -|\text{Alice, O}z,\, -) = 1 \tag{21.149}$$

- quand il se trompe d'axe, il ne trouve (statistiquement) le même résultat qu'Alice que dans la moitié des cas (voir (21.139)) :

$$P(\text{Bob, O}x,\, +|\text{Alice, O}z,\, +) = \frac{1}{2} \;,\; P(\text{Bob, O}x,\, -|\text{Alice, O}z,\, +) = \frac{1}{2} \;. \tag{21.150}$$

$$P(\text{Bob, O}x,\, +|\text{Alice, O}z,\, -) = \frac{1}{2} \;,\; P(\text{Bob, O}x,\, -|\text{Alice, O}z,\, -) = \frac{1}{2} \;. \tag{21.151}$$

et de même en permutant x et z. Compte tenu de la séquence d'Alice (21.147), un exemple de séquence pour Bob, *en l'absence d'espion*, est le suivant, la deuxième ligne donnant les choix d'axes de Bob, la troisième étant les résultats de ses mesures :

$$\begin{array}{ccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ \updownarrow & \leftrightarrow & \updownarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \leftrightarrow & \updownarrow & \updownarrow & \leftrightarrow & \updownarrow & \leftrightarrow & \updownarrow \\ + & + & - & - & + & + & + & + & - & - & + & + & - & + & + \end{array} \; ; \tag{21.152}$$

l'absence d'espion est signée par le fait que quand la direction de mesure est la même entre (21.147) et (21.152), le résultat est *forcément* le même. Dans le cas contraire, Bob peut néanmoins trouver le même résultat qu'Alice.

Toutes les mesures étant faites, Alice fait part publiquement[67] à Bob de tous ses *choix* d'axes, mais, pour l'instant, ne dit évidemment rien de ses *résultats* ; réciproquement, Bob lui communique ses axes. Ceci leur permet de faire le tri, et d'éliminer tous les cas où le hasard a fait que les axes n'étaient pas les mêmes pour Alice et pour Bob :

n° de la mesure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Alice	$\updownarrow$	$\updownarrow$	$\leftrightarrow$	$\updownarrow$	$\leftrightarrow$	$\leftrightarrow$	$\leftrightarrow$	$\updownarrow$	$\leftrightarrow$	$\leftrightarrow$	$\updownarrow$	$\leftrightarrow$	$\updownarrow$	$\updownarrow$	$\leftrightarrow$
Bob	$\updownarrow$	$\leftrightarrow$	$\updownarrow$	$\updownarrow$	$\leftrightarrow$	$\leftrightarrow$	$\updownarrow$	$\leftrightarrow$	$\leftrightarrow$	$\updownarrow$	$\updownarrow$	$\leftrightarrow$	$\updownarrow$	$\leftrightarrow$	$\updownarrow$
Retenus	+			−	+	+			−		+	+	−		

(21.153)

À ce stade (et toujours en l'absence d'espion), chacun sait que pour tous les *bits* ainsi retenus, il a dû trouver les mêmes résultats que l'autre[68] ; toutefois, en l'état actuel des choses, nul ne sait ce qu'il en est au juste (si c'est faux, un espion écoute !).

Pour le savoir, Alice communique maintenant à Bob *une partie* de ses résultats (elle les *sacrifie*) ; connaissant les choix d'axes et les résultats d'Alice pour un assez grand nombre de cas, un test statistique permet alors à Bob de savoir avec une faible probabilité d'erreur si la communication a été écoutée ou non. Si la confidentialité est alors établie, la séquence des *autres* résultats – non-communiqués, donc à l'abri inviolable de toute écoute ! – constituent la clé codant une lettre convenue d'avance : *sans rien communiquer d'autre*, Alice et Bob savent qu'ils ont trouvé la même chose à chaque fois puisqu'ils savent avoir utilisé les mêmes axes. Finalement, comme la clé est constituée sans l'échange de rien d'autre de plus, aucun moyen n'existe d'altérer la confidentialité de la transmission de celle-ci !

Il faut bien voir que l'espion éventuel ne peut pas extraire quoi que ce soit d'utile : dans la phase d'envoi, les spins ne véhiculent aucune information en eux-mêmes. En fait, l'information ne devient réalité qu'après la mise en œuvre du scénario décrit ci-dessus.

Une variante consiste à utiliser une source produisant des *paires* (a, b) de spins $S = 1/2$ dans l'état singulet de leur somme[69] $\vec{S} = \vec{S}_\mathrm{a} + \vec{S}_\mathrm{b}$:

$$|0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) , \tag{21.154}$$

développé sur les états propres de $S_{z\,\mathrm{a}}$ et $S_{z\,\mathrm{b}}$. Puis, les particules sont séparées, l'une allant vers Alice, l'autre filant vers Bob ; par la suite, les deux observateurs procèdent chacun à des mesures consécutives de la composante le long de Ox ou de Oz du spin qu'ils recoivent, directions dont ils sont convenus d'avance (ce sont par exemple les deux directions perpendiculaires à la direction de propagation des particules). On aura aussi

[67] On entend par là que la communication ne nécessite aucune confidentialité.

[68] Chacun sait ce qu'a trouvé l'autre sans y être allé voir et sans communication explicite !

[69] En faisant la somme de deux spins 1/2, on peut construire un état singulet ($S = 0$, une composante, $M = 0$), et un état triplet ($S = 1$, trois composantes $M = 0, \pm 1$).

besoin ci-dessous des états propres $\pm\frac{\hbar}{2}$ de S_x, notés $|\rightarrow\rangle$ et $|\leftarrow\rangle$ respectivement ; compte tenu de :

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \ , \qquad |\leftarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle) \ , \tag{21.155}$$

l'état singulet (21.154) se décompose comme suit :

$$|0\rangle = \frac{1}{\sqrt{2}}(|\leftarrow\rightarrow\rangle - |\rightarrow\leftarrow\rangle) \ . \tag{21.156}$$

Soit d'abord la séquence où Alice et Bob mesurent le long de la même direction, Oz par exemple. Préséances obligent, Alice effectue la première une mesure, de S_z par exemple ; si elle trouve $+\frac{\hbar}{2}$, l'état des deux particules après cette mesure est la projection de (21.154), soit $|\uparrow\downarrow\rangle$. Il en résulte que quand Bob va mesurer par la suite la composante S_z de *son* spin, sa mesure porte sur l'état réduit par la mesure antérieure d'Alice, soit $|\uparrow\downarrow\rangle$, et il trouvera $-\frac{\hbar}{2}$ avec certitude[70]. Au contraire, si Alice trouve $-\frac{\hbar}{2}$, Bob trouvera plus tard $+\frac{\hbar}{2}$: dans tous les cas, leurs résultats sont donc à coup sûr opposés. Il en va de même si ce sont des mesures de S_x qui sont effectuées : le même raisonnenment vaut avec (21.156). En définitive, quand Alice et Bob mesurent la même composante, leurs résultats sont toujours *anticorrélés*. Physiquement et en résumé, sachant que l'état est singulet et qu'Alice a trouvé le spin dans un sens, Bob ne peut trouver que le spin opposé.

Si Alice mesure S_z et Bob S_x, les choses vont autrement. Si Alice a trouvé $+\frac{\hbar}{2}$, l'état après cette mesure est $|\uparrow\downarrow\rangle$ et se décompose suivant :

$$\frac{1}{2}(|\rightarrow\rightarrow\rangle - |\rightarrow\leftarrow\rangle + |\leftarrow\rightarrow\rangle - |\leftarrow\leftarrow\rangle) \ . \tag{21.157}$$

Bob trouvera donc pour S_x les deux valeurs $\pm\frac{\hbar}{2}$ avec équiprobabilité. Il en va de même si Alice trouve $-\frac{\hbar}{2}$.

Ainsi, dans tous les cas, chacun peut savoir (*a priori* sans communication des résultats : c'est le miracle EPR) ce que l'autre a trouvé, tout ceci bien sûr en l'absence d'espion. Ce savoir *potentiel* se concrétise par la communication mutuelle des axes ; dès lors, le *sacrifice* d'une partie des résultats permet, tout comme dans la première version, de tester la confidentialité de la ligne.

♦ *Remarque*

La fiabilité du jugement d'Alice et Bob sur leur certitude de ne pas être écoutés repose sur le fait que toute interception de leurs spins donne lieu presque toujours à une *altération* de l'état de ceux-ci. On pourrait aussi imaginer que l'espion est assez astucieux pour se borner à regarder passer les spins sans rien faire d'autre qu'une

[70] La mesure permet ainsi d'obtenir instantanément des informations sur ce qui se passe ailleurs : c'est ici que réside le paradoxe EPR [126] (à ce sujet, on pourra consulter l'article de Jammer [127] et de Laloë [128]) – voir aussi section 21.5.

copie de leur état, mais sans mesurer quoi que ce soit (il se borne à les *photographier* au passage). Ceci est en réalité impossible, en vertu du théorème de non-clonage quantique [129], [130], dont la démonstration peut être résumée comme suit.

Soit un certain système A dans l'état[71] $|\Psi\rangle_A$, dont on souhaite faire une copie. Pour cela, on considère un autre système, b, situé dans l'état initial $|\psi\rangle_b$, qui est l'état "blanc" sur lequel on veut essayer de copier $|\Psi\rangle_A$. Bien évidemment, l'état $|\psi\rangle_b$ ne saurait dépendre de $|\Psi\rangle_A$, dont on ne sait rien *a priori*. Une première façon de faire est de laisser interagir ces deux systèmes d'une façon ou d'une autre, puis d'effectuer une mesure d'observable : clairement, ceci va corrompre l'état $|\Psi\rangle_A$, et ne peut donner qu'une copie dégradée, voire dénaturée. Ce n'est clairement pas ce que l'on souhaite.

La seule et unique autre possibilité est de laisser évoluer le système de façon unitaire, par son opérateur d'évolution $U(t)$; au bout d'un certain temps, le vecteur d'état est devenu $|\Psi(t)\rangle = U(t)|\Psi\rangle_A \otimes |\psi\rangle_b$. La copie *neutre* (clonage) aura été effectuée si l'état $|\Psi(t)\rangle$ est égal à $|\Psi\rangle_A \otimes |\Psi\rangle_b$: l'état de A est inaltéré, et le système b se retrouve dans l'état où était (et est encore) le système A. Il faut donc :

$$U(t)|\Psi\rangle_A \otimes |\psi\rangle_b = |\Psi\rangle_A \otimes |\Psi\rangle_b \ . \tag{21.158}$$

Bien évidemment, ceci doit être vrai quel que soit l'état initial du système A ; la copie est donc possible si, quels que soient deux états $|\Psi\rangle_A$ et $|\Phi\rangle_A$, on a l'égalité (21.158) et :

$$U(t)|\Phi\rangle_A \otimes |\psi\rangle_b = |\Phi\rangle_A \otimes |\Phi\rangle_b \ . \tag{21.159}$$

Comme l'opérateur U est unitaire, il conserve le produit scalaire : deux états quelconques transformés par U, et les mêmes états avant transformation, ont le même produit scalaire. On a donc :

$$(|\Psi\rangle_A \otimes |\Psi\rangle_b,\ |\Phi\rangle_A \otimes |\Phi\rangle_b) = (|\Psi\rangle_A \otimes |\psi\rangle_b,\ |\Phi\rangle_A \otimes |\psi\rangle_b) \tag{21.160}$$

Le premier membre est égal à $(|\Psi\rangle_A,\ |\Phi\rangle_A) \times (|\Psi\rangle_b,\ |\Phi\rangle_b)$, tandis que le second vaut $(|\Psi\rangle_A,\ |\Phi\rangle_A) \times (|\psi\rangle_b,\ |\psi\rangle_b)$; il vient donc :

$$\boxed{\langle\Psi|\Phi\rangle^2 = \langle\Psi|\Phi\rangle} \tag{21.161}$$

Cette égalité n'est possible que dans deux cas : $|\Phi\rangle = |\Psi\rangle$ *ou* $|\Phi\rangle \perp |\Psi\rangle$, ce qui n'est pas vrai en général, puisque $|\Phi\rangle$ doit pouvoir être choisi n'importe comment par rapport à $|\Psi\rangle$. Un opérateur unitaire ne peut donc pas copier (*cloner*) un état quantique.

Ce théorème est donc une conséquence directe de l'unitarité de l'évolution, et implique de ce fait la *linéarité* foncière de la théorie quantique. Pour mettre ceci en évidence, prenons précisément le cas où $|\Psi\rangle$ et $|\Phi\rangle$ sont orthogonaux, auquel cas, l'égalité (21.161) est satisfaite, et soit à effectuer une copie de l'état général $\lambda|\Phi\rangle + \mu|\Psi\rangle$, λ et μ quelconques reliés par $|\lambda|^2 + |\mu|^2 = 1$; pour que ceci soit possible, il faut que l'évolution suivante le soit aussi :

$$(\lambda|\Phi\rangle + \mu|\Psi\rangle) \otimes |\psi\rangle_b \stackrel{???}{\longrightarrow} (\lambda|\Phi\rangle + \mu|\Psi\rangle) \otimes (\lambda|\Phi\rangle + \mu|\Psi\rangle) \tag{21.162}$$

[71] Tous les états considérés sont supposés normalisés à l'unité.

soit :

$$(\lambda|\Phi\rangle + \mu|\Psi\rangle) \otimes |\psi\rangle_b \overset{???}{\longrightarrow} \lambda^2|\Phi\rangle \otimes |\Phi\rangle + \lambda\mu(|\Phi\rangle \otimes |\Psi\rangle + |\Psi\rangle \otimes |\Phi\rangle) + \mu^2|\Psi\rangle \otimes |\Psi\rangle \ . \tag{21.163}$$

Le procédé de copiage fonctionne si on a pour chaque état "de base" $|\Phi\rangle$ et $|\Psi\rangle$:

$$|\Phi\rangle \otimes |\psi\rangle_b \longrightarrow |\Phi\rangle \otimes |\Phi\rangle \ , \qquad |\Psi\rangle \otimes |\psi\rangle_b \longrightarrow |\Psi\rangle \otimes |\Psi\rangle \ ; \tag{21.164}$$

en vertu de la linéarité de l'équation de Schrödinger, ceci entraîne :

$$(\lambda|\Phi\rangle + \mu|\Psi\rangle) \otimes |\psi\rangle_b \longrightarrow \lambda|\Phi\rangle \otimes |\Phi\rangle + \mu|\Psi\rangle \otimes |\Psi\rangle \ . \tag{21.165}$$

C'est donc cet état qui est obtenu en partant de $(\lambda|\Phi\rangle + \mu|\Psi\rangle) \otimes |\psi\rangle_b$: l'état final (21.163) est donc inaccessible, et il est différent de l'état réel (21.165), sauf bien sûr si le produit $\lambda\mu$ est nul. La linéarité est ainsi clairement responsable de l'impossibilité de cloner un état quantique en général.

L'existence de ce théorème est donc cruciale pour la cryptographie quantique. À l'inverse, il interdit *a priori* la possibilité de procéder à des vérifications lors d'un calcul utilisant les principes de la Mécanique quantique : pendant quelque temps, ceci a été considéré comme une limitation fatale pour l'ordinateur quantique (sans parler d'autres [131]), car la mise en jeu de codes correcteurs d'erreur est une nécessité dans tout procédé de calcul automatique. En réalité, il a été possible de concevoir des codes de correction *quantiques* qui permettent de s'accommoder du théorème de non-clonage. ♦

21.4 Décohérence

La notion de cohérence quantique, selon la définition de Cohen-Tannoudji, a été introduite p. 926. La cohérence est responsable des oscillations (aux fréquences de Bohr) des valeurs moyennes d'une observable qui n'est pas une constante du mouvement, quand l'état du système est non-stationnaire. Elle apparaît alors en tant qu'éléments non-diagonaux $\rho_{nm}(t)$ de la matrice de l'opérateur densité $\rho(t)$ représenté sur la base propre du Hamiltonien, en particulier lorsque $\rho(t)$ représente un cas pur ; dans ce cas, $\rho_{nm}(t)$ est le produit de deux facteurs (voir (21.35)), $\rho_{nm}(t) = c_n c_m^* \mathrm{e}^{-\mathrm{i}(\omega_n - \omega_m)t}$.

Le phénomène de *décohérence* désigne généralement le fait que, pour une raison ou pour une autre, les éléments non-diagonaux de $\rho(t)$ sur une base donnée ne sont pas constants en module et décroissent au cours du temps. On a vu que lorsque toutes les cohérences sont *nulles*, ρ représente un cas mixte extrême[72] pourvu que deux populations au moins sont différentes de zéro. La décohérence est donc l'un des mécanismes par lesquels un cas pur évolue en cas mixte. Fondamentalement, la décohérence peut se produire dès que l'on prend en compte l'*environnement* d'un système[73] : c'est dire qu'elle survient

[72] Il n'est pas nécessaire que les cohérences soient nulles pour avoir un cas mixte.

[73] L'environnement d'un système est en quelque sorte l'antithèse du *milieu subquantique* proposé par de Broglie [132].

toujours, un système n'étant jamais isolé, que ce soit vis-à-vis d'un appareil de mesure, ou du reste du monde. Ajoutons d'ailleurs que le mot *environnement* ne désigne l'extérieur du système que si celui-ci est *microscopique*. Pour un système macroscopique, on peut définir une (ou des) variable(s) classique(s) (par exemple le centre de gravité d'une boule de billard), dont l'*environnement* est constitué, pour partie, de degrés de liberté internes au sens le plus strict : les phonons représentant les vibrations (quantiques !) des $N \ggg 1$ atomes de la boule sont de parfaits candidats pour former cet environnement nécessaire, et d'ailleurs trivialement inévitable. En définitive, la question fondamentale semble pouvoir se réduire à la construction de modèles aussi universels que possible décrivant l'effet de l'environnement d'un système, et d'examiner le *temps de décohérence* qui définit l'échelle au-delà de laquelle les éléments non-diagonaux sont pratiquement nuls.

Au plan conceptuel, la décohérence joue un rôle majeur puisque, selon qu'elle survient ou pas, les interférences quantiques s'évanouissent ou, au contraire, demeurent. De surcroît, comme on l'a discuté p. 928, la modélisation d'une mesure idéale consiste, partant d'un cas pur, à annuler *à la main* toutes les cohérences, ce qui revient à singer pragmatiquement une décohérence totale et instantanée et traduit, en termes de l'opérateur densité, la réduction du paquet d'ondes provoquée par une mesure d'observable. De toute évidence, le phénomène de décohérence est aussi au centre de la question du passage graduel (?) du régime quantique au régime classique, et doit permettre de comprendre le comportement classique observé pour les objets macroscopiques : on verra de fait que, dans des conditions physiques raisonnables (notamment de température), le temps de décohérence d'un système macroscopique est si court qu'il n'a pas réellement de sens physique, sauf à être considéré comme nul[74]. Enfin, la décohérence est au cœur du problème de l'ordinateur quantique, puisque les capacités phénoménales de ce dernier reposent crucialement sur la préservation de la cohérence de phase d'une superposition d'états quantiques, et la manipulation de cette dernière. L'une des raisons situant aujourd'hui cet ordinateur en Utopie est précisément l'incapacité actuelle d'assurer une telle préservation. Si l'importance de la notion de décohérence est réalisée depuis longtemps (voir par exemple la discussion profonde qu'en donne Bohm [134], chapitre 22, notamment p. 602), elle a été revivifiée par Zeh [135]. L'engouement suscité par les attraits de l'ordinateur quantique a amplifié le regain d'intérêt pour cette question, déjà fondamentale en soi vis-à-vis de la théorie de la mesure[75].

De façon quelque peu schématique, le phénomène de décohérence se traduit techniquement pour les éléments non-diagonaux[76] par une variation temporelle du type :

$$|\rho_{nm}(t)| = |\rho_{nm}(0)| f_{nm}(t/\tau_{nm}) \ , \tag{21.166}$$

où f_{nm} est une certaine fonction positive décroissante, et τ_{nm} une échelle de temps

[74] Omnès ([133], p. 94) utilise cette formule :

"... *la cause principale de la position de la Lune (principale, car il y en a d'autres) tient simplement au fait que le Soleil l'éclaire.*"

[75] Pour un point récent sur l'ensemble de ces questions, voir [136].

[76] Rappelons que la *matrice* de $\rho(t)$ acquiert un sens physique direct dès que l'on a fait le choix de l'observable objet d'une mesure, ou d'une analyse de la transition classique $\longrightarrow$ quantique –, voir le commentaire suivant l'équation (21.48).

caractéristique. La toute première question à considérer est de préciser, au moins qualitativement, comment un tel comportement peut survenir. Une analyse intuitive permet alors d'accepter définitivement l'idée suivant laquelle, fondamentalement, la décohérence résulte du couplage de tout système avec son environnement, lequel introduit inéluctablement le *hasard* quand il s'agit de décrire un système réel. La référence à la notion classique de cohérence optique permet notamment de réaliser qu'une accumulation de phases aléatoires conduit inévitablement à une loi du genre (21.166) ; l'argument suivant est certes semi-quantitatif, mais sa valeur n'en est pas moins à la fois exemplaire et universelle – sa simulation avec une calculette permet de le voir à l'œuvre.

Soit $\mathrm{e}^{\mathrm{i}\alpha(t)}$, où $\alpha(t)$ est une phase résultant de l'addition de N petits incréments aléatoires $\delta\alpha_i \equiv \delta\alpha(t_i)$ survenant à des instants t_i entre $t = 0$ et t :

$$\alpha(t) = \sum_i \delta\alpha_i \ . \tag{21.167}$$

Sans perdre de généralité, on peut supposer que les $\delta\alpha_i$ ont une moyenne nulle, et qu'ils sont décorrrélés au bout d'un certain temps fini Δt. En divisant le segment $[0,\, t]$ selon une grille *moins fine* que Δt, on obtient une nouvelle somme constituant une *moyenne à gros grain*[77] $\alpha(t) = \sum_I \delta\alpha_I$, où cette fois, chaque α_I (lui-même somme d'accroissements plus élémentaires) est décorrélé de tous les autres ; la somme donnant $\alpha(t)$ contient maintenant $N \sim \frac{t}{\Delta t}$ termes. En supposant que les accroissements de phase aléatoires $\delta\alpha_I$ ont un écart-type *fini* $\sigma_{\delta\alpha}^2$, on peut invoquer le théorème limite central ([100], [137]) : la phase $\alpha(t)$, étant une somme en grand nombre ($N \ggg 1$) de variables centrées décorrélées de variance finie, est asymptotiquement une variable *gaussienne* centrée de variance $N\sigma_{\delta\alpha}^2$:

$$\langle \alpha^2 \rangle \simeq N\sigma_{\delta\alpha}^2 \sim \frac{\sigma_{\delta\alpha}^2}{\Delta t}\, t \ . \tag{21.168}$$

Pour une variable gaussienne centrée X d'écart-type σ_X, la fonction caractéristique est $\phi(k) \stackrel{\text{déf}}{=} \langle \mathrm{e}^{\mathrm{i}kX} \rangle = \mathrm{e}^{-k^2\sigma_X^2}$; on a donc ici (asymptotiquement) :

$$\boxed{\langle \mathrm{e}^{\mathrm{i}\alpha} \rangle = \mathrm{e}^{-\gamma t} \ , \qquad \gamma \stackrel{\text{déf}}{=} \frac{\sigma_{\delta\alpha}^2}{\Delta t}} \tag{21.169}$$

Ceci montre que la moyenne des facteurs de phase décroît (ici) comme une exponentielle : une accumulation de phases aléatoires est donc de nature à produire un comportement tel que (21.166), au moins en moyenne. Une analyse plus élaborée est possible (voir par exemple [138]) – où l'effet du hasard sur les phases conduit à une équation maîtresse pour l'opérateur densité –, mais produit essentiellement le même résultat, d'ailleurs fort banal à qui est familier avec le mouvement Brownien[78], ou la cohérence en Optique.

Ce comportement est caractéristique d'une évolution *irréversible* (relaxation), qui ne peut d'aucune façon résulter de l'équation de Liouville pour l'opérateur densité global

[77] une opération généralement appelée *coarse-graining* en anglais.

[78] Le problème du mouvement Brownien a été réglé grâce aux contributions magistrales de Perrin [139] et Einstein [140] au tout début du 20e siècle. Il est au cœur de la théorie quantique en raison de la proximité technique de l'intégrale de Feynman et celle de Wiener [141].

d'un système isolé : comme on l'a vu, l'équation de Liouville (21.38) correspond à une évolution *unitaire*, qui ne peut transformer un cas pur en cas mixte[79]. Au contraire, lorsqu'un sous-système est couplé à un autre, l'opérateur densité *réduit* du sous-système peut effectivement changer de nature : pur au départ, il peut fort bien évoluer en cas mixte. Le prototype d'une telle situation est constitué d'un système d'intérêt couplé à un grand système, ce dernier jouant le rôle d'un *réservoir*, encore appelé *bain*, et mimant aussi bien que possible l'environnement de tout système réel – environnement au sens large, comme précisé p. 963. Une question majeure en pareille circonstance est la définition d'un modèle où le bain possède une certaine universalité, tout en autorisant des calculs traitables. Le modèle de Caldeira et Leggett[80] ([103], [146]) est précisément une tentative particulièrement féconde dans cette direction, mettant notamment en évidence une brisure de symétrie constituant un scénario possible pour la décohérence. Ce résultat majeur, obtenu au sein d'une approximation[81] (*dilute-blip approximation*) dans le formalisme de Feynman - Vernon [104], a été controversé mais semble devoir conserver l'essentiel de sa valeur.

Il a été déjà expliqué pourquoi la *matrice* de l'opérateur densité acquiert un sens physique direct quand une observable donnée a été choisie pour l'analyse. Jusqu'à présent, seules des représentations *discrètes* ont été considérées, la base propre du Hamiltonien par exemple[82]. On peut aussi décider de focaliser sur une observable continue, la position q d'une particule par exemple, auquel cas l'opérateur densité ρ est représenté[83] par un *noyau* (matrice continue) $\rho(q,\, q'\, ;\, t)$; pour un cas pur associé à la fonction d'onde $\Psi(q,\, t)$, le noyau de ρ est :

$$\boxed{\rho(q,\, q'\, ;\, t) \equiv \langle q|\rho(t)|q'\rangle = \langle q|\Psi(t)\rangle\langle\Psi(t)|q'\rangle = \Psi(q,\, t)\Psi^*(q',\, t)} \tag{21.170}$$

Les éléments non-diagonaux $\rho(q,\, q',\, t)$ sont alors les *cohérences spatiales* (voir problème 21.6.4), cependant que les éléments diagonaux donnent la (densité de) probabilité de trouver la particule autour du point d'abscisse q lors d'une mesure de position.

[79]Selon Zurek [142] :

"Unitary evolution condemns every closed system to "purity"*."*

[80]Anthony James Leggett a partagé le Nobel en 2003 avec Alexei A. Abrikosov et Vitaly L. Ginzburg *"for pioneering contributions to the theory of superconductors and superfluids"*. L'une des motivations du modèle spin - boson était précisément l'analyse de la brisure de symétrie d'un spin fortement couplé à un environnement ; Caldeira et Leggett ont mis en lumière un mécanisme par lequel une particule (quantique !) dans un double puits et couplée à un bain, au lieu d'effectuer des oscillations amorties d'un côté à l'autre par effet tunnel, reste partiellement piégée là où elle était au départ, lorque le couplage à l'environnement dépasse une certaine valeur-seuil finie.

Un mécanisme du même genre a été proposé par Primas ([143], [144]) et Pfeifer [145] il y a assez longtemps, pour expliquer l'existence d'isomères optiques (énantiomères) – qui constituent un bel exemple de symétrie brisée ; en pareil cas, le bain est constitué par le champ électromagnétique (quantifié !).

[81]Cette approximation est équivalente à une resommation partielle obtenue dans le cadre plus conventionnel d'une approximation de Born [147].

[82]En toute rigueur, les états propres non-liés du Hamiltonien – quand ils existent – doivent être introduits, donnant, même pour H, une représentation hybride ayant une composante discrète *et* une composante continue.

[83]Pour la légéreté des notations, on considère *une* seule particule dans $\mathbb{R}$; l'extension formelle à N particules dans $\mathbb{R}^D$ est un simple jeu d'écriture.

Lorsqu'il s'agit d'analyser la transition quantique → classique, le noyau $\rho(q,\, q'\,;\, t)$ pose une question de fond : dans le cadre quantique élémentaire, on peut choisir la représentation-q *ou* la représentation-p, ces deux écritures étant exclusives puisque définies vis-à-vis de deux observables incompatibles ($[q,\, p] \neq 0$!). Au contraire, dans le cadre classique, l'espace de référence est *l'espace des phases* : la dynamique d'un système classique étant déterminée complètement par la valeur à l'instant t de la position $q(t)$ et du moment conjugué $p(t)$, l'état du système est sans aucune ambiguïté représenté par un certain point de coordonnées $(q(t),\, p(t))$ dans l'espace des phases, qui est ici $\mathbb{R}^2$. Au fil du temps, ce point décrit une trajectoire (orbite) ; pour un système hamiltonien (pas de dissipation), cette orbite ne peut se recouper que si le mouvement est périodique – et alors elle se *superpose* à elle-même périodiquement.

Observé avec un éclairage quantique, l'espace des phases ne peut être considéré comme un ensemble de *points* puisque, par les relations de Heisenberg, la plus petite cellule envisageable est un domaine dont l'extension est *finie*, respectant $\Delta p\Delta x \gtrsim \hbar$: à l'échelle quantique, l'espace des phases est un ensemble de "gros points". Le choix d'une représentation, -q *ou* -p, revient à raisonner en considérant cet espace comme une juxtaposition de plans épais perpendiculaires à l'axe associé à l'observable fournissant la base de représentation : choisir de définir la densité de probabilité de la coordonnée q rejette l'impulsion, en tant que variable, dans la plus totale indétermination. Il est manifeste que cette partition asymétrique de l'espace ne se prête guère à l'analyse de la transition graduelle[84] (?) vers le domaine classique, que l'on imagine plutôt comme le résultat de la diminution effective de la taille des "points" de volume fini[85] d'ordre $\hbar$.

Dans les circonstances présentes, le noyau $\rho(q,\, q'\,;\, t)$ ne semble donc pas adéquat, *a priori*. En revanche, Wigner [148] a introduit une représentation mixte de l'opérateur densité, notée $W(q,\, p,\, t)$, qui semble au contraire parfaitement appropriée ; pour un cas pur associé à une fonction d'onde $\Psi(q,\, t)$, sa définition est la suivante :

$$\boxed{W(q,\, p,\, t) \stackrel{\text{déf}}{=} \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p\xi}\, \Psi(q-\frac{\xi}{2},\, t)\Psi^*(q+\frac{\xi}{2},\, t)\, \mathrm{d}\xi} \tag{21.171}$$

La fonction de Wigner est donc essentiellement une transformation de Fourier impliquant une sorte de convolution, qui *explore* symétriquement l'espace autour du point q fixé, et sonde en fait les régions où le paquet d'ondes décrit par $\Psi(q,\, t)$ a une valeur notable. L'intégration de $W(q,\, p,\, t)$ sur l'une des variables q ou p donne la densité de probabilité de présence dans la représentation de l'autre variable ; par exemple, compte tenu de $\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k = 2\pi\delta(x)$ et de $\delta(ax) = \frac{1}{|a|}\delta(x)$:

$$\int_{-\infty}^{+\infty} W(q,\, p,\, t)\,\mathrm{d}p = \frac{1}{2\pi}\int_{-\infty}^{+\infty} 2\pi\delta(\xi)\Psi(q-\frac{\xi}{2},\, t)\Psi^*(q+\frac{\xi}{2},\, t)\,\mathrm{d}\xi = |\Psi(q,\, t)|^2\ . \tag{21.172}$$

[84] La *douceur* de cette transition n'a rien d'évident : tout au contraire, les exemples se sont multipliés montrant comment la limite classique est hautement singulière.

[85] On rencontre ici un bel exemple de l'importance des *échelles* en Physique : pour un système macroscopique, dont l'action typique est d'ordre $10^n\hbar$, avec $n \ggg 1$ – l'action S_{Terre} de la Terre dans sa révolution sidérale est $\sim 10^{74}\hbar$ (voir Tome I, éq. (9.59), note 26 p. 608, section 16.5, etc.) –, le domaine de dimension linéaire $\sim \hbar$ peut être considéré comme un infiniment petit, autorisant à considérer l'espace comme un ensemble de points et effaçant *ipso facto* toute trace de quantification. À l'inverse, pour un système quantique, la *granulation* de l'espace des phases est incontournable.

On peut ainsi utiliser la fonction de Wigner pour calculer la valeur moyenne d'une fonction quelconque de q seul, $A(q)$:

$$\langle A(q)\rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W(q,\,p,\,t)A(q)\,\mathrm{d}q\mathrm{d}p \tag{21.173}$$

Par ailleurs, l'intégrale $\int_{\mathbb{R}} W(q,\,p,\,t)\,\mathrm{d}q$ est la transformée de Fourier de la convolution $\Psi \star \Psi^*$: par le théorème de convolution, c'est donc le produit $\Phi(p)\Phi^*(p)$, où $\Phi(p)$ est la transformée de Fourier de $\Psi(q)$, c'est-à-dire la représentation-p de l'état en question :

$$\int_{-\infty}^{+\infty} W(q,\,p,\,t)\,\mathrm{d}q = |\Phi(p,\,t)|^2\ , \qquad \Phi(p) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}pq}\,\Psi(q,\,t)\,\mathrm{d}q\ . \tag{21.174}$$

On a donc tout autant, pour une observable ne dépendant que de p :

$$\langle B(p)\rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W(q,\,p,\,t)B(p)\,\mathrm{d}q\mathrm{d}p \tag{21.175}$$

Toutefois, la fonction $W(q,\,p,\,t)$ n'est pas une *densité*, car elle peut prendre des valeurs négatives, ce qui est bien naturel puisque, telle qu'elle est définie, cette fonction est spécifiquement quantique et incorpore donc tous les effets possibles d'interférences ; d'un autre côté, elle est normalisée au sens où $\int_{\mathbb{R}}\mathrm{d}q\int_{\mathbb{R}}\mathrm{d}p\,W(q,\,p,\,t) = 1$. Par ailleurs, $W(q,\,p,\,t)$ permet bien de reconstruire $\rho(q,\,q'\,;\,t)$; par exemple la relation inverse de (21.171) s'écrit :

$$\rho(q,\,q'\,;\,t) \equiv \Psi(q,\,t)\Psi^*(q',\,t) = \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(q'-q)p}\,W\big(\frac{q+q'}{2},\,p\,,\,t\big)\,\mathrm{d}p \tag{21.176}$$

À titre d'illustration, pour un paquet d'ondes pris gaussien par simplicité (voir problème 21.6.3) :

$$\Psi(q) \stackrel{\text{déf}}{=} \frac{1}{\sqrt{\sqrt{2\pi}\Delta q}}\,\mathrm{e}^{-\frac{q^2}{4\Delta q^2}}\ , \tag{21.177}$$

on trouve :

$$\boxed{W(q,\,p) = \frac{1}{\pi\hbar}\,\mathrm{e}^{-\frac{q^2}{2\Delta q^2}}\,\mathrm{e}^{-\frac{2\Delta q^2}{\hbar^2}p^2}} \tag{21.178}$$

Visiblement, la fonction de Wigner est la représentation quantique de l'opérateur densité la plus similaire à la description classique : elle fournit ici une gaussienne dont les deux écarts-types sont précisément reliés solidairement suivant la relation de Heisenberg : s'il est parfaitement loisible de concevoir l'espace des phases comme un ensemble de points au sens ordinaire, aucune particule quantique n'échappe à la dispersion inévitable de la coordonnée et de l'impulsion.

La définition de la fonction de Wigner se généralise sans peine pour un cas mixte ; pour un opérateur densité construit avec des fonctions (normalisées) $\Psi_n(q,\,t)$:

$$\rho(q,\,q'\,;\,t) = \sum_n p_n\,\Psi_n(q,\,t)\Psi_n^*(q',\,t) \qquad (p_n \geq 0\ ,\ \sum_n p_n = 1)\ , \tag{21.179}$$

on écrit par extension de (21.171) :

$$\boxed{W(q,\,p,\,t) \stackrel{\text{déf}}{=} \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p\xi}\,\rho(q-\frac{\xi}{2},\,q+\frac{\xi}{2},\,t)\,\mathrm{d}\xi} \tag{21.180}$$

définition qui tient dans tous les cas, pur ou mixte.

Comme l'explique Zurek [142], l'adéquation de $W(q,\,p,\,t)$ est magnifiquement illustrée par la considération, soit d'un cas mixte formé avec une combinaison (à poids égaux) de deux gaussiennes normalisées $\Psi_{\pm}(q) \stackrel{\text{déf}}{=} \Psi(q \mp \frac{l}{2})$ situées à la distance l l'une de l'autre, soit par un cas pur dont la fonction d'onde associée est une combinaison linéaire des deux mêmes gaussiennes. Dans le premier cas, on a :

$$W_{\text{mixte}}(q,\,p) = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p\xi}\,\frac{1}{2}\sum_{\sigma=\pm}\Psi_{\sigma}(q+\frac{\xi}{2})\Psi_{\sigma}^{*}(q-\frac{\xi}{2})\,\mathrm{d}\xi \ . \tag{21.181}$$

Avec la fonction gaussienne (21.177), on trouve :

$$\boxed{W_{\text{mixte}}(q,\,p) = \frac{1}{2\pi\hbar}\sum_{\sigma=\pm 1}\mathrm{e}^{-\frac{(q-\sigma l/2)^2}{2\Delta q^2}}\,\mathrm{e}^{-\frac{2\Delta q^2 p^2}{\hbar^2}}} \tag{21.182}$$

Si $\Delta q \ll l$, ceci représente deux paquets "classiques" séparés de la distance l, et bien isolés : vers $q \simeq 0$, cette fonction est exponentiellement petite en raison de $\Delta q \ll l$ (fig. 21.10, à gauche).

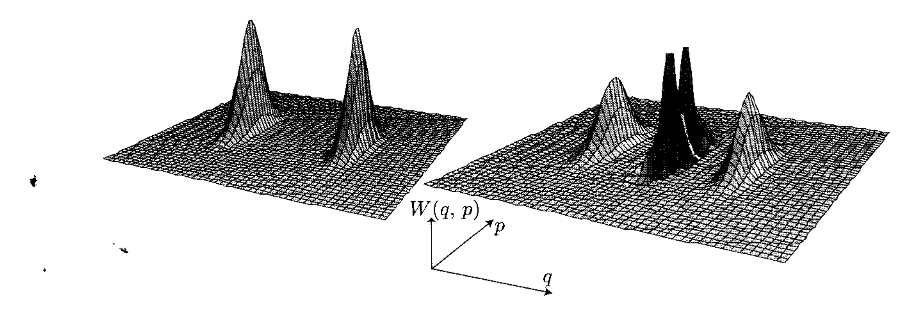


Figure 21.10: Fonction de Wigner pour le cas mixte (21.181) (à gauche) et pour le cas pur construit avec la superposition cohérente (21.183) (à droite). Les interférences centrales sont nettement visibles (voir l'expression approchée (21.185)).

Soit maintenant la superposition cohérente $|\Psi_{\text{coh}}\rangle \stackrel{\text{déf}}{=} [2(1+S)]^{-1/2}[|\Psi_{+}\rangle + |\Psi_{-}\rangle]$, qui est une sorte de *Chat de Schrödinger* ; le facteur $S(l) \stackrel{\text{déf}}{=} \langle\Psi_{-}|\Psi_{+}\rangle$ prend en compte le fait que le produit scalaire des deux fonctions $\Psi_{\pm}(q)$ n'est pas nul (il vaut $\mathrm{e}^{-l^2/(8\Delta q^2)}$). Avec cette fonction $\Psi_{\text{coh}}(q)$, on peut former le cas pur $\rho_{\text{pur}}(q,\,q') = \Psi_{\text{coh}}(q)\Psi_{\text{coh}}^{*}(q')$, et aussi la fonction de Wigner :

$$W_{\text{pur}}(q,\,p) = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p\xi}\,\Psi_{\text{coh}}(q-\frac{\xi}{2})\Psi_{\text{coh}}^{*}(q+\frac{\xi}{2})\,\mathrm{d}\xi \ ; \tag{21.183}$$

on trouve facilement :

$$W_{\text{pur}}(q,\,p) = \frac{1}{1+S(l)}\left[W_{\text{mixte}}(q,\,p) + \frac{1}{\pi\hbar}\mathrm{e}^{-\frac{q^2}{2\Delta q^2}}\,\mathrm{e}^{-\frac{2\Delta q^2}{\hbar^2}p^2}\cos\frac{lp}{\hbar}\right]\ . \tag{21.184}$$

Avec l'hypothèse $\Delta q \ll l$, il vient[86] :

$$\boxed{W_{\text{pur}}(q,\,p) \simeq W_{\text{mixte}}(q,\,p) + \frac{1}{\pi\hbar}\mathrm{e}^{-\frac{q^2}{2\Delta q^2}}\,\mathrm{e}^{-\frac{2\Delta q^2}{\hbar^2}p^2}\cos\frac{lp}{\hbar}} \tag{21.185}$$

Cette expression montre d'une part que la fonction de Wigner peut effectivement être d'un signe variable : près de $q = 0$, le premier terme est exponentiellement petit, et seul compte le second. D'autre part, elle met bien en évidence la différence fondamentale entre un cas pur et un cas mixte, construits chacun avec strictement les mêmes ingrédients : pour le cas pur, apparaît un terme d'interférences, ici de nature spatiale, avec une longueur d'onde (variable avec le plan de coupe) égale à $\frac{h}{p}$ (voir fig. 21.10 à droite) ; on note aussi que, toujours près du plan $q = 0$, $W_{\text{pur}}(q,\,p)$ prend des valeurs notables, alors que $W_{\text{mixte}}(q,\,p)$ est quasi-nul : pour un cas pur *à deux bosses* bien distinctes, il existe des zones où la probabilité est à peu près nulle de trouver la particule là où pourtant la fonction de Wigner est loin d'être nulle.

Il apparaît ainsi clairement que la fonction de Wigner se prête particulièrement bien à l'examen des questions où la décohérence joue un rôle central, qu'il s'agisse de la transition quantique → classique, ou de la théorie de la mesure. En définitive, Wigner définit un cadre unifié se référant à l'espace des phases où les deux mondes quantique et classique trouvent leur place, caractérisés par des régimes non seulement également représentables, mais aussi identifiables sans aucune ambiguïté.

À ce stade de l'analyse, la question naturelle est la suivante : peut-on imaginer un processus dynamique par lequel un cas pur tel que (21.183) évoluerait vers le cas mixte (21.181) ? On doit être convaincu que cette irréversibilité nécessaire ne peut survenir qu'en faisant interagir le système d'intérêt avec son environnement (le bain), lequel est la source d'une forme d'aléatoire imposant une flèche au temps. On peut alors tenter de répondre à cette question de deux façons au moins :

- soit en visant d'obtenir l'opérateur densité réduit pour le système d'intérêt, permettant de trouver la valeur moyenne de quelques observables et de chercher la signature du passage classique → quantique dans la relaxation *anormale* (?) de certaines d'entre elles. C'est au fond la direction prise par Caldeira et Leggett.
- soit en raisonnant d'emblée avec la fonction de distribution de Wigner, et en essayant de construire une équation du type Fokker - Planck incorporant au mieux les processus sous-jacents responsables de la perte de cohérence au cours du temps ; cette approche a été suivie notamment par Walls et Milburn [138] et par Brune et al. [149].

[86] L'expression (21.185), approchée, est normalisée à des termes $\mathcal{O}(\mathrm{e}^{-l^2/\Delta q^2}) \ll 1$ près.

Par analogie avec le mouvement Brownien, la première manière ressemble à un traitement *à la Langevin* : on écrit une équation dynamique, et on essaie d'en extraire une part d'information (vitesse moyenne, fluctuations de vitesse, etc.). La seconde vise à obtenir le maximum d'information escomptable dans un cadre probabiliste, en ayant pour objectif l'obtention des lois de probabilités elles-mêmes.

Dans la deuxième démarche, l'obtention d'une équation d'évolution n'est pas chose aisée, mais peut être conduite à terme dans quelques cas, par exemple pour une particule massive de masse M dans un bain de particules légères ([102], [150][87]) – ce qui correspond très exactement à la situation caractérisant le mouvement Brownien classique. On trouve ainsi, pour une particule d'énergie potentielle $V(q)$ dans un bain d'oscillateurs[88] ([142], [151]) :

$$\frac{\partial W}{\partial t} = -\frac{p}{M}\frac{\partial W}{\partial q} + \frac{\partial V}{\partial q}\frac{\partial W}{\partial p} + \gamma\frac{\partial}{\partial p}(pW) + D\frac{\partial^2 W}{\partial p^2} \ , \tag{21.186}$$

où γ est l'inverse d'un temps, et où D est un coefficient de diffusion pour le moment conjugué p. L'interprétation de cette équation d'apparence compliquée est en fait très aisée. Les deux premiers termes ne sont rien d'autre que le crochet de Poisson $\{H,\, W\}$ pour un Hamiltonien égal à $\frac{p^2}{2M} + V(q)$; à eux seuls, ils représentent exclusivement les termes formellement classiques, obtenus dans un cadre purement mécanique où aucune forme d'aléatoire n'est présente. Les deux termes suivants sont d'une tout autre nature : ils recopient à l'identique la divergence d'un certain courant $j(p,\, t)$ pour évoquer le second membre d'une équation de conservation générique $\partial_t P(x,\, t) = -\partial_x j(x,\, t)$ traduisant l'écoulement sans pertes d'un fluide (de probabilité par exemple), de densité (concentration) $P(x,\, t)$ – mais ne pas oublier que $W(q,\, p,\, t)$ n'est pas une densité. Ici, le courant $j(q,\, p,\, t) \stackrel{\text{déf}}{=} -\gamma pW - D\partial_p W$ contient les deux composantes habituelles : un terme de dérive traduisant un frottement, et un terme de diffusion proportionnel au gradient de "concentration" (nouvelle apparition d'un avatar de la loi de Fick) ; cette dérive et cette diffusion concernent exclusivement l'impulsion de la particule et traduisent – tout comme dans le mouvement Brownien standard – l'action des "collisions" entre la particule massive et les particules légères du bain, seule source d'aléatoire et responsable de la diffusion du moment conjugué. L'interprétation de γ comme un coefficient de frottement se confirme en multipliant membre à membre (21.186) par p et en intégrant sur q et p. Le premier membre donne $\frac{\mathrm{d}}{\mathrm{d}t}\langle p\rangle$; après une intégration par parties, le terme $\gamma\frac{\partial}{\partial p}(pW)$ devient $-\gamma\langle p\rangle$, supposant que $\lim_{|p|\to\infty}(p^2 W) = 0$; les autres termes sont aussi nuls (refaire des intégrations par parties), et il reste :

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle p\rangle = -\gamma\langle p\rangle \ . \tag{21.187}$$

De même, le coefficient D apparaît bien comme mesurant la diffusion de l'impulsion ; en multipliant membre cette fois par p^2, on trouve après intégration sur q et p :

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle p^2\rangle = -2\gamma\langle p^2\rangle + 2D \ . \tag{21.188}$$

[87] Ces auteurs fournissent aussi de nombreux exemples, assortis d'estimations de la valeur numérique du temps de décohérence. Faut-il vraiment prendre au sérieux l'application de la théorie de la décohérence au cerveau humain ? (voir sous-section 3.2.5 de cet ouvrage).

[88] C'est à nouveau le modèle de Caldeira et Leggett qui est sous-jacent.

La variation de $\langle p^2 \rangle$ est donc la combinaison d'un amortissement dû à la friction, et d'un accroissement à taux constant égal à $2D$.

Une équation telle que (21.186) est aussi obtenue pour une particule lourde dans un bain de particules légères [102], en calculant les incréments $\delta\vec{p}$ et $\delta\vec{p}^2$ de l'impulsion $\vec{p}$ de la particule par suite des collisions. En effectuant à nouveau une analyse *à gros grain* en temps, on trouve (encore une conséquence du théorème limite central) que $\delta\vec{p}$ et $\delta\vec{p}^2$ sont tous deux proportionnels à l'intervalle de temps δt. Ces procédés sont exactement les mêmes que ceux utilisés pour l'équation de Fokker - Planck à propos du mouvement Brownien classique et, tout naturellement, produisent une équation du même type en ce qui concerne les termes d'origine stochastique.

Cela étant dit, (21.186) se récrit de façon plus transparente :

$$\boxed{\frac{\partial W}{\partial t} = \{H,\, W\} - \frac{\partial}{\partial p} j(q,\, p,\, t) \ , \qquad j(q,\, p,\, t) = -\gamma p W - D\frac{\partial W}{\partial p}} \tag{21.189}$$

Ainsi, le second membre rassemble toutes les contributions à la variation de $W(q,\, p,\, t)$ en un point fixé de l'espace des phases : celles qui décrivent le mouvement purement mécanique (hamiltonien) et qui coïncident formellement strictement avec les termes classiques, et une contribution venant de la décohérence induite par une diffusion d'impulsion par suite des collisions. Au total, cette équation réalise une sorte de fusion assez extraordinaire[89] entre une évolution purement mécanique, et une équation de type Fokker - Planck – quoique la fonction de Wigner n'est pas une densité de probabilité[90]. $W(q,\, p,\, t)$ possède certaines apparences d'une grandeur classique, mais reste fondamentalement un objet quantique, en conséquence directe de sa définition même.

On peut alors montrer, au moins dans des cas simples, que l'équation (21.189) assure bien la transformation d'un cas pur en un cas mixte ; l'échelle de temps d'une telle *mutation* est l'élément crucial à regarder de près, car c'est elle qui va conditionner dans quelle mesure on peut espérer observer l'évolution quantique $\rightarrow$ classique, par exemple, ou encore, de façon plus radicale, expliquer pourquoi des boules de billard ne relèvent pas de la Mécanique quantique, étant entendu que pour les observer, il faut bien *éclairer* le tapis vert, au moins. Le point essentiel est d'examiner l'allure du terme de relaxation dans des situations exemplaires ; pour des particules massives, on trouve que ce temps caractéristique $\tau_{\text{décoh}}$ est tout naturellement construit avec les seules échelles disponibles :

$$\boxed{\tau_{\text{décoh}} \sim \gamma^{-1}\, \frac{\hbar^2}{Ma^2 k_{\text{B}} T}} \tag{21.190}$$

a étant une échelle de longueur pertinente pour la situation analysée[91]. La température intervient en conséquence du fait que le bain est supposé être à l'équilibre thermodynamique, ce qui contraint d'ailleurs les deux constantes D et γ à satisfaire la relation

[89]L'équation (21.189) est une sorte d'*équation de Boltzmann* [99] où cohabitent des termes purement hamiltoniens et une *intégrale de collisions*.

[90]L'apparition d'un courant exactement au même titre que pour une vraie densité de probabilité vient du fait que l'intégrale dans tout l'espace de phases de $W(q,\, p,\, t)$ est une constante.

[91]À la réflexion, il peut sembler que l'expression du produit $\gamma\tau_{\text{décoh}}$ aurait pu être écrite sans le moindre calcul, juste en manipulant les échelles d'énergie disponibles. En fait, il n'existe pas vraiment

d'Einstein $\frac{D}{M\gamma} = k_{\mathrm{B}}T$. Pour une particule de 1 g, pour $a \sim 1$ cm et $T = 300$ K, on a $\gamma\tau_{\text{décoh}} \sim \times 10^{-41}$: comme le fait observer Zurek [142], quand bien même le temps de relaxation γ^{-1} serait égal à ... l'âge de l'univers, soit $\sim 10^{17}$ s, le temps de décohérence serait de l'ordre de 10^{-24} s ; pour que l'on ait $\tau_{\text{décoh}} \sim \gamma^{-1}$, pour cette même particule, il faudrait que la température ne soit guère supérieure à... 10^{-38} K. À l'inverse, il a été possible de suivre le processus de décohérence dans des systèmes *mésoscopiques* [152].

La théorie de la décohérence fournit ainsi un schéma explicatif permettant de réconcilier les deux points de vue, classique et quantique. Le couplage inévitable de tout système à son environnement provoque une disparition graduelle des cohérences, et transforme inéluctablement tout cas pur en cas mixte. La grandeur - clé est le temps de décohérence : pour des systèmes mésoscopiques, il est suffisamment long pour pouvoir être observé [152] ; à l'inverse, pour un système macroscopique, il est si petit que sa valeur numérique n'a d'autre de sens que de devoir être considérée comme un infiniment petit, au sens des échelles physiques pertinentes, rejoignant ainsi le clan de ces grandeurs physiques considérées comme nulles, soit d'un point de vue pragmatique car en-deçà des capacités expérimentales, soit sur le plan conceptuel (la masse du photon, la charge du neutron, l'inverse de la durée de vie du proton...). En particulier, la réduction *instantanée* du paquet d'ondes énoncée en tant que postulat ne doit pas être considérée comme incorrectement décrite, étant donné l'extrême petitesse du temps de décohérence pour un système macroscopique[92].

Cette théorie ne répond pas à toutes les questions de fond ; par exemple, elle n'explique pas comment *une valeur* est choisie plutôt qu'une autre lors d'une mesure d'observable. Elle fait le silence sur un hypothétique mécanisme de *sélection*, dont l'existence semble indiscutable si l'on adopte une attitude déterministe pure et dure. Elle n'explique pas non plus l'*unicité* du monde macroscopique (le Chat est-il mort ou vivant ?). Il reste que la théorie de la décohérence a reçu à ce jour de nombreuses confirmations expérimentales.

21.5 Intrication

L'intrication a pour conséquence des phénomènes parmi les plus étranges de tous ceux prévus par la Mécanique quantique, et oblige notamment à réviser radicalement la notion de *séparabilité* de deux systèmes. On admet implicitement en Physique classique que si deux systèmes sont très éloignés l'un de l'autre, la dynamique propre de chacun et les observations faites sur l'un d'entre eux ne dépendent en aucune façon de l'autre, et n'ont

d'argument convaincant permettant d'affirmer *a priori* que ce paramètre ne doit dépendre que du rapport des deux énergies caractéristiques pertinentes. En outre, la théorie permet d'estimer raisonnablement le coefficient de frottement γ.

[92]Il en va de même pour le Second principe de la Thermodynamique : le temps de Poincaré d'un système macroscopique est si grand que le déclarer par principe *infini* constitue une simple affirmation de bon sens.

aucune action sur l'autre ; deux tels systèmes sont dits séparés, encore que cette appellation, dans le cadre classique, ne semble pas d'une grande utilité tellement elle apparaît triviale.

En Mécanique quantique, il en va tout autrement, et ce fut le point de départ d'Einstein, Podolsky et Rosen dans leur célèbre analyse [126] tendant à prouver que, malgré les immenses succès déjà acquis, la théorie quantique était une théorie physique *incomplète.* C'est un fait *quantique* que deux systèmes ayant interagi ne sont plus *séparables,* même s'ils se retrouvent finalement aussi loin que l'on veut l'un de l'autre : la séparation *spatiale* ne signifie plus la *séparation* au sens où tout ce qui arrive à l'un d'entre eux n'a aucun effet en retour sur l'autre.

La non-séparabilité de deux systèmes ayant interagi est une conséquence directe du formalisme quantique, et peut se comprendre très simplement. De toute évidence, l'argument suivant correspond à une expérience idéale, et il convient de le relativiser le cas échéant en tenant compte, par exemple, des phénomènes de décohérence. Il est donné pour introduire l'idée... et aussi parce que l'expérience idéale en question a été accomplie par Aspect et al. ([153], [154]) ; ses résultats[93] donnent tort à Einstein, Podolsky et Rosen, et confirment pleinement ce qu'affirme la Mécanique quantique, aussi extravagant que cela paraisse à première vue.

Considérons deux systèmes, dénotés A et B, pour lesquels il est possible, en l'absence de toute interaction mutuelle, de définir des états quantiques, propres d'un ECOC contenant par exemple le Hamiltonien de chacun d'entre eux ; on peut ainsi écrire :

$$\exists\ \{|\psi_{Ak}\rangle\}_k,\ \{|\psi_{Bl}\rangle\}_l\ :\ H_A|\psi_{Ak}\rangle = E_{Ak}|\psi_{Ak}\rangle\ ,\ \ H_B|\psi_{Bl}\rangle = E_{Bl}|\psi_{Bl}\rangle\ . \tag{21.191}$$

Dans l'état initial, ces systèmes sont supposés à l'infini l'un de l'autre ; par la linéarité de l'équation de Schrödinger, on ne restreint pas la généralité en supposant que chacun d'entre eux est dans l'un des états propres, $|\psi_{Ak_\mathrm{i}}\rangle$ pour l'un, $|\psi_{Bl_\mathrm{i}}\rangle$ pour l'autre ; avec l'hypothèse de séparation initiale, l'état primitif du système composite $(A+B)$ est donc le produit tensoriel :

$$|\Psi_\text{initial}\rangle = |\psi_{Ak_\mathrm{i}}\rangle \otimes |\psi_{Bl_\mathrm{i}}\rangle \equiv |\psi_{Ak_\mathrm{i}},\ \psi_{Bl_\mathrm{i}}\rangle\ . \tag{21.192}$$

Supposons maintenant que les deux systèmes se trouvent provisoirement en interaction – ce sont par exemple deux particules issues de deux sources très éloignées, qui entrent en collision puis repartent chacune *à l'infini.* L'interaction fait évoluer le vecteur d'état de façon éventuellement compliquée, et produit finalement un vecteur d'état possédant *a priori* (toute règle de sélection mise à part) des composantes sur tous les états du genre $|\psi_{Ak}\rangle \otimes |\psi_{Bl}\rangle$; formellement :

$$|\Psi_\text{final}\rangle \equiv U(t_\text{final},\ t_\text{initial})|\Psi_\text{initial}\rangle = \sum_k \sum_l c_{kl}|\psi_{Ak},\ \psi_{Bl}\rangle\ . \tag{21.193}$$

[93]L'expérience d'Aspect et al. compte parmi les plus belles expériences de la Physique contemporaine. Son caractère *quasi-idéal* peut être globalement apprécié en comparant la valeur mesurée de l'indice de Bell défini en (21.202), soit $S_\text{Bell, exp} = 2,697 \pm 0,015$, remarquablement proche de sa valeur maximale théorique $S_\text{Bell, max} = 2\sqrt{2} = 2,828...$ (voir (21.206)).

Les coefficients c_{kl} ne sont rien d'autre que les amplitudes de transition :

$$\langle \psi_{Ak},\, \psi_{Bl} | U(t_{\text{final}},\, t_{\text{initial}}) | \psi_{Ak_i},\, \psi_{Bl_i} \rangle \ , \tag{21.194}$$

en principe calculables si l'on connaît l'interaction. Clairement, l'expression (21.193) est une conséquence immédiate indiscutable avec les hypothèses acceptées, et ne présente à première vue rien d'extraordinaire.

En réalité, la décomposition (21.193) révèle à la réflexion une propriété pour le moins surprenante : dans l'état final, où les deux systèmes sont séparés physiquement au sens le plus ordinaire du terme, on ne peut plus attribuer à chacun un *état* bien défini : au contraire de l'état initial *factorisé* (21.192), ou chacun est *de son côté*, l'état final est une superposition *emmêlée* de tous les couples d'états possibles. Un tel état est dit[94] *intriqué*. L'exemple sans doute le plus simple d'état de cette nature est la composante $|S = 1,\, M_S = 0\rangle$ de l'état triplet pour deux spins 1/2, puisqu'elle s'écrit :

$$|S = 1,\, M_S = 0\rangle = \frac{1}{\sqrt{2}}\Big(| +\frac{1}{2}\rangle \otimes | -\frac{1}{2}\rangle + | -\frac{1}{2}\rangle \otimes | +\frac{1}{2}\rangle \Big) \ ; \tag{21.195}$$

il en va d'ailleurs de même pour le singulet, dont l'expression est :

$$|S = 0,\, M_S = 0\rangle = \frac{1}{\sqrt{2}}\Big(- | +\frac{1}{2}\rangle \otimes | -\frac{1}{2}\rangle + | -\frac{1}{2}\rangle \otimes | +\frac{1}{2}\rangle \Big) \ . \tag{21.196}$$

La manipulation d'états intriqués simples permet de mettre en évidence la plus fameuse des *inégalités de Bell* [155], [156]. Considérons à nouveau deux spins 1/2, désignés A et B (comme Alice et Bob), appelés *qubits* dans le jargon à la mode. Notons[95] simplement $|0\rangle$ et $|1\rangle$ les états propres de la matrice de Pauli σ_z :

$$\sigma_z|0\rangle = |0\rangle \ , \qquad \sigma_z|1\rangle = -|1\rangle \ . \tag{21.197}$$

Introduisons maintenant les quatre états[96] dits *de Bell* :

$$|\Phi^{(\pm)}\rangle \overset{\text{déf}}{=} \frac{1}{\sqrt{2}}(|0,\, 0\rangle \pm |1,\, 1\rangle) \ , \qquad |\Psi^{(\pm)}\rangle \overset{\text{déf}}{=} \frac{1}{\sqrt{2}}(|0,\, 1\rangle \pm |1,\, 0\rangle) \ , \tag{21.198}$$

où les labels $|q_1,\, q_2\rangle$ se rapportent toujours à A et B respectivement. De toute évidence, il s'agit d'états intriqués – remarquer que les $|\Psi^{(\pm)}\rangle$ ne sont rien d'autre que les composantes $M_S = 0$ des deux états singulet et triplet (21.195) et (21.196) –, qui de surcroît se déduisent les uns des autres par des opérations très simples :

$$\sigma_z^A|\Phi^{(\pm)}\rangle = \sigma_z^B|\Phi^{(\pm)}\rangle = |\Phi^{(\mp)}\rangle \ , \qquad \sigma_z^A|\Psi^{(\pm)}\rangle = -\sigma_z^B|\Psi^{(\pm)}\rangle = |\Psi^{(\mp)}\rangle \ , \tag{21.199}$$

$$\pm\sigma_x^A|\Phi^{(\pm)}\rangle = \sigma_x^B|\Phi^{(\pm)}\rangle = |\Psi^{(\pm)}\rangle \ , \qquad \pm\sigma_x^A|\Psi^{(\pm)}\rangle = \sigma_x^B|\Psi^{(\pm)}\rangle = |\Phi^{(\pm)}\rangle \ . \tag{21.200}$$

[94] En anglais *entangled.* Schrödinger fut semble-t-il le premier à employer le mot *Vershrankung.*

[95] La notation a aussi le bon goût (?) d'avoir un petit air d'informatique...

[96] Ces états jouent également un rôle important pour la *téléportation quantique* [157], protocole permettant, avec notamment la complicité d'une troisième particule, de transférer à distance l'information incluse dans un état quantique.

Soit maintenant la composante σ_Z obtenue de σ_z par une rotation de θ autour de Oy, égale[97] à $\sigma_Z = \sigma_z \cos\theta + \sigma_x \sin\theta$; les relations (21.199) et (21.200) entraînent que si l'on mesure cette composante sur chacun des deux spins dans l'état $|\Phi^+\rangle$, on trouve toujours la valeur +1 ; en effet :

$$\sigma_Z^A \sigma_Z^B |\Phi^{(+)}\rangle = [\cos^2\theta\sigma_z^A\sigma_z^B + \sin^2\theta\sigma_x^A\sigma_x^B + \sin\theta\cos\theta(\sigma_z^A\sigma_x^B + \sigma_x^A\sigma_z^B)]|\Phi^{(+)}\rangle = |\Phi^{(+)}\rangle \ . \tag{21.201}$$

Une fois préparé un tel état intriqué, les deux spins de la paire sont séparés spatialement, l'un étant récupéré par Alice, l'autre par Bob. La coïncidence des résultats d'une mesure exprimée par (21.201) permet à Alice de savoir la valeur de son spin sans procéder à une mesure : il suffit qu'elle demande à Bob de faire, lui, une mesure selon une certaine direction θ qu'elle lui indique. Ainsi, Alice peut connaître son spin sans l'avoir mesuré, donc sans l'avoir perturbé en quoi que ce soit ; de ce fait, le bon sens semble imposer l'idée que la valeur du spin d'Alice préexiste à toute opération de mesure... qu'Alice n'a pas faite. Comme on peut inverser les rôles, le même bon sens semble forcer à conclure que les valeurs des deux spins préexistent à l'opération de mesure ; on dit que c'est une *variable cachée* : elle a une valeur en tant que telle, que l'on découvre cette valeur par une mesure ou qu'on la laisse dans l'ombre faute d'avoir effectué la moindre mesure[98]. Raisonner ainsi, c'est admettre que de telles variables existent, dont on précise d'ailleurs qu'elles sont *locales*, pour exprimer le fait qu'elles existent en un lieu donné, indépendamment de ce qui peut se passer ailleurs.

Maintenant, Alice et Bob envisagent d'effectuer des mesures, chacun de son spin, sans se concerter sur la direction ; par exemple Alice choisira selon son humeur Ox ou Oz, Bob de son côté choisira OX ou OZ, tournés de θ par rapport aux axes d'Alice. Formons maintenant la quantité :

$$S_{\text{Bell}} \overset{\text{déf}}{=} (p_z^A s_z^A - p_x^A s_x^A) p_Z^B s_Z^B + (p_z^A s_z^A + p_x^A s_x^A) p_X^B s_X^B \tag{21.202}$$

où p_z^A est la probabilité qu'Alice tire la direction Oz, etc., les s_z^A,... étant tous des nombres ne pouvant prendre que l'une des deux valeurs +1 ou −1. Cette expression peut aussi se récrire en notant $\langle ... \rangle$ les valeurs moyennes (au sens des espérances mathématiques) :

$$S_{\text{Bell}} \overset{\text{déf}}{=} \langle \sigma_z^A \sigma_Z^B \rangle - \langle \sigma_x^A \sigma_Z^B \rangle + \langle \sigma_z^A \sigma_X^B \rangle + \langle \sigma_x^A \sigma_X^B \rangle \tag{21.203}$$

Dans l'expression (21.202), chacun des facteurs $(p_z^A s_z^A \pm p_x^A s_x^A)$ est borné en module par $|p_z^A - p_x^A|$ et $p_z^A + p_x^A = 1$, tandis que les deux autres facteurs sont, toujours en module, compris entre 0 et 1. Il en résulte immédiatement que :

$$\boxed{|S_{\text{Bell}}| \leq 2} \tag{21.204}$$

En conclusion, supposer que les spins ont une certaine valeur (pas *certaine* !) indépendamment de toute mesure, effectuée ou non[99], borne *supérieurement* l'indice de Bell.

[97] La composante perpendiculaire est $\sigma_X = -\sigma_z \sin\theta + \sigma_x \cos\theta$.

[98] Une théorie à variables cachées devrait prendre en compte le rôle de l'appareil de mesure, faute de quoi on peut éliminer la possibilité de telles variables [158].

[99] L'argument de Bell n'exclut pas la possibilité d'une action perturbative de l'appareil de mesure : la valeur effectivement mesurée – quand elle l'est – peut encore différer de la valeur supposée préexistante.

Le point remarquable est qu'il existe des orientations θ qui *violent* l'inégalité (21.204). Prenons $\theta = \pi/4$; la valeur moyenne $\langle\sigma_z^A\sigma_Z^B\rangle$ vaut :

$$\langle\Phi^{(+)}|\sigma_z^A\sigma_Z^B|\Phi^{(+)}\rangle = \frac{1}{\sqrt{2}}\langle\Phi^{(+)}|\sigma_z^A(\sigma_z^B + \sigma_x^B)|\Phi^{(+)}\rangle \;; \tag{21.205}$$

compte tenu des égalités (21.199) et (21.200), l'action de l'opérateur sur le *ket* donne $|\Phi^{(+)}\rangle + |\Psi^{(-)}\rangle$, d'où $\langle\sigma_z^A\sigma_Z^B\rangle = +\frac{1}{\sqrt{2}}$; de la même façon, on trouve $\langle\sigma_z^A\sigma_X^B\rangle = -\frac{1}{\sqrt{2}}$, $\langle\sigma_x^A\sigma_X^B\rangle = +\frac{1}{\sqrt{2}}$ et $\langle\sigma_x^A\sigma_Z^B\rangle = +\frac{1}{\sqrt{2}}$. Reportant dans (21.203), on trouve :

$$S_{\text{Bell}} = +\frac{1}{\sqrt{2}} - (-\frac{1}{\sqrt{2}}) + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = 4\frac{1}{\sqrt{2}} = 2\sqrt{2} > 2 \tag{21.206}$$

Pour un angle θ quelconque, on trouve $S_{\text{Bell}} = 2\sqrt{2}\sin(\theta + \frac{\pi}{4})$: l'égalité de Bell est donc violée si $0 < \theta < \frac{\pi}{2}$ (π). C'est cette violation qui a été prouvée expérimentalement par Aspect et al. ([153], [154]) en analysant les corrélations entre deux photons issus d'une même transition atomique, les deux photons étant détectés aux bouts des deux bras opposés d'un interféromètre situés à plusieurs mètres de distance l'un de l'autre. Depuis cette époque, de nombreuses autres expériences ont été conduites ; leurs résultats concordants permettent d'affirmer que l'on dispose aujourd'hui d'une accumulation de preuves expérimentales de la violation des inégalités de Bell [159].

La conclusion est que la logique classique, qui consiste à raisonner avec des valeurs *non mesurées* supposées existantes, est incompatible avec les lois de la Mécanique quantique, ce que l'on résume parfois par la formule suivante en forme de truisme *"Une mesure non faite n'a pas de résultat"*. Le résultat de l'expérience d'Aspect et al. fournit la démonstration de l'inexistence de variables cachées locales, auxquelles il pouvait être légitime de croire, et qui ont été l'objet de formalisations très élaborées pendant des décennies.

Terminons cette section en revenant sur l'étrangeté majeure projetée sur le devant de la scène par Einstein, Podolsky et Rosen, en rediscutant l'expression de l'état final (21.193) des deux particules une fois séparées dans l'espace. L'aspect paradoxal de la situation éclate quand on imagine que l'on effectue la mesure d'une certaine observable sur l'*un* des deux systèmes ; afin de fixer les idées, admettons que l'on mesure l'énergie de A, et qu'il n'existe aucune dégénérescence. Les postulats affirment alors que l'on va trouver une certaine énergie E_{Ak_f} appartenant à l'ensemble $\{E_{Ak}\}_k$, avec une certaine probabilité – tout autant calculable en principe. C'est ici qu'apparaît le paradoxe : juste après cette mesure, le système $(A + B)$ est dans l'état réduit $|\Psi_{\text{final}}\rangle_{\text{juste après}}$ tel que le système A est dans l'état E_{Ak_f} (la valeur trouvée par la mesure sur A) cependant que, par la conservation de l'énergie (système isolé !), le système B se trouve forcément dans l'état d'énergie parfaitement déterminée E_{Bl_f}, satisfaisant l'égalité $E_{Ak_\text{f}} + E_{Bl_\text{f}} = E_{Ak_\text{i}} + E_{Bl_\text{i}}$:

$$|\Psi_{\text{final}}\rangle_{\text{juste après}} = |\psi_{Ak_\text{f}},\, \psi_{Bl_\text{f}}\rangle \;. \tag{21.207}$$

La bizarrerie n'est pas de connaître immédiatement l'énergie de B quand on a trouvé celle de A : il en irait de même classiquement[100]. Ce qui est pour le moins surprenant, c'est

[100] Quand on connaît l'énergie initiale de deux boules de billard, et que l'on sait la collision élastique, la mesure de l'énergie d'une boule donne immédiatement l'énergie de l'autre boule – *remake* du jeu du voyageur et du sédentaire, voir Tome I, note 27, p. 399.

que la mesure effectuée sur A modifie instantanément *l'état* de B, qui est à l'autre bout du monde. Autrement dit, effectuer une mesure sur A – dont le résultat est d'ailleurs aléatoire – provoque une action à distance instantanée sur le système B, même si celui-ci se trouve à des années-lumière : c'est là le paradoxe mis en avant par le trio EPR, dont la conclusion était d'affirmer que, suivant *leur* définition de la *réalité physique*, la Mécanique quantique est une théorie *incomplète*. Aujourd'hui, l'infirmation expérimentale directe de cette conclusion se complète en déclarant que la Mécanique quantique est une théorie intrinsèquement *non-locale*[101]. Dans sa formulation traditionnelle, la Mécanique quantique peut s'accommoder de la violation de la localité au prix de l'indétermination des grandeurs physiques. Pour une discussion profonde de ce caractère particulier, la lecture des articles pénétrants de Mermin [160], [161] est vivement recommandée, tout comme la conclusion de l'article d'Aspect et Grangier [159].

♦ *Remarque*

Ces dernières années, il s'est avéré que les atomes de Rydberg sont des systèmes se prêtant remarquablement à la mise en évidence des effets quantiques spécifiques, qu'il s'agisse de manipuler des *chats de Schrödinger*, des états intriqués ou d'observer en direct la fonction de Wigner – pour une revue, voir [162] et [163]. ♦

21.6 Exercices et problèmes

21.6.1 Traitement phénoménologique d'un atome à trois niveaux

Soit un atome à trois niveaux, $|\mathrm{g}\rangle$, $|\mathrm{m}\rangle$ et $|\mathrm{e}\rangle$; $|\mathrm{g}\rangle$ est l'état fondamental, $|\mathrm{m}\rangle$ un état excité métastable dont la désexcitation radiative vers le fondamental a une probabilité très faible, que l'on néglige dans la suite. $|\mathrm{e}\rangle$ est un état fortement couplé au fondamental par échange de photons avec l'extérieur ; pour fixer les idées, l'ordre des énergies est $E_\mathrm{g} < E_\mathrm{m} < E_\mathrm{e}$. La fuite radiative à partir de $|\mathrm{e}\rangle$ vers le fondamental est représentée *effectivement* par une partie imaginaire $-\mathrm{i}\hbar\Gamma_\mathrm{e}/2$ ajoutée formellement à l'énergie de l'état $|\mathrm{e}\rangle$ ($\Gamma_\mathrm{e} > 0$) ; $\tau_\mathrm{e} \stackrel{\text{déf}}{=} \Gamma_\mathrm{e}^{-1}$ est la durée de vie purement radiative de l'état $|\mathrm{e}\rangle$. Cette introduction effective est la façon la plus élémentaire d'éliminer les autres degrés de liberté et constitue l'image la plus triviale de la sommation (trace partielle), dans des situations plus complexes, sur des degrés de liberté n'agissant plus qu'en coulisse. Dans la suite, on considère donc exclusivement les deux niveaux $|\mathrm{m}\rangle$ et $|\mathrm{e}\rangle$, supposés satisfaire les équations suivantes :

$$H_\mathrm{at}|\mathrm{m}\rangle = \hbar\omega_\mathrm{m}|\mathrm{m}\rangle \qquad H_\mathrm{at}|\mathrm{e}\rangle = \hbar(\omega_\mathrm{e} - \mathrm{i}\frac{\Gamma_\mathrm{e}}{2})|\mathrm{e}\rangle \ . \tag{21.208}$$

[101] Le qualificatif *non-local* est souvent associé à la possibilité (?) de transmission instantanée par des signaux supra-luminiques. Ici, le terme est pris au sens le plus strict du terme : ce qui se passe à un endroit est corrélé avec ce qui se passe ailleurs.

L'atome, initialement dans l'état $|\mathrm{m}\rangle$, est perturbé par un champ externe (semi-classique) harmonique résonnant représenté par l'opérateur[102] $V(t)$:

$$V(t) = \hbar\Omega \left[|\mathrm{e}\rangle\langle\mathrm{m}|\, \mathrm{e}^{\mathrm{i}(\omega_\mathrm{m}-\omega_\mathrm{e})t} + |\mathrm{m}\rangle\langle\mathrm{e}|\, \mathrm{e}^{\mathrm{i}(\omega_\mathrm{e}-\omega_\mathrm{m})t}\right] \quad . \tag{21.209}$$

Le vecteur d'état dans le secteur $(|\mathrm{m}\rangle, |\mathrm{e}\rangle)$ admet la décomposition :

$$|\Psi(t)\rangle = a_\mathrm{m}(t)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{m}t}|\mathrm{m}\rangle + a_\mathrm{e}(t)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{e}t}|\mathrm{e}\rangle \tag{21.210}$$

On peut convenir que la largeur Γ_e est fixée (physiquement, c'est une caractéristique intrinsèque de l'état excité) ; Ω mesurant le couplage avec le champ extérieur, $\lambda \stackrel{\text{déf}}{=} \Omega/\Gamma_\mathrm{e}$ est finalement le seul paramètre du problème.

1. Écrire l'équation de Schrödinger (effective) pour les coefficients $a_\mathrm{m}(t)$ et $a_\mathrm{e}(t)$.
2. Comment varient $a_\mathrm{m}(t)$ et $a_\mathrm{e}(t)$ au tout début du mouvement ($t \ll \Omega^{-1}$, Γ_e^{-1}).
3. Que représente la différence $1 - (|a_\mathrm{m}(t)|^2 + |a_\mathrm{e}(t)|^2)$ $\forall\, t$?
4. Utiliser la transformation de Laplace pour intégrer le système.
5. Dans la situation de couplage fort ($\lambda \gg 1$) :
   (a) Écrire les expressions approchées de $a_\mathrm{m}(t)$ et $a_\mathrm{e}(t)$, et des probabilités $P_\mathrm{m}(t)$ et $P_\mathrm{e}(t)$ de trouver respectivement l'atome dans les états $|\mathrm{m}\rangle$ et $|\mathrm{e}\rangle$.
   (b) À quelle fréquence s'effectue l'oscillation ?
   (c) Combien vaut approximativement la somme $|a_\mathrm{m}(t)|^2 + |a_\mathrm{e}(t)|^2$? Tracer sa variation en fonction du temps ; observer l'existence de petits paliers.
6. Dans la situation de couplage faible ($\lambda \ll 1$) :
   (a) Écrire les expressions approchées de $a_\mathrm{m}(t)$ et $a_\mathrm{e}(t)$, et des probabilités $P_\mathrm{m}(t)$ et $P_\mathrm{m}(t)$.
   (b) Comment varie cette fois la population de l'état fondamental $|\mathrm{g}\rangle$?
   (c) Identifier la durée de vie effective de l'état $|\mathrm{m}\rangle$.
7. Quel est le sens physique de $\tau_{\mathrm{eff,\,m}} \stackrel{\text{déf}}{=} \int_0^{+\infty} P_\mathrm{m}(t)\,\mathrm{d}t$? Trouver l'expression de $\tau_{\mathrm{eff,\,m}}$ en donner les formes approchées dans les deux cas $\lambda \ll 1$ et $\lambda \gg 1$.

[102] Le champ extérieur donnant lieu au couplage $V(t)$ avec l'atome est *résonnant*, puisqu'il oscille précisément à la pulsation de Bohr $\omega_\mathrm{e} - \omega_\mathrm{m}$ de l'atome.

21.6.2 Effet Zénon sur un neutron

Le neutron est un fermion de spin $S = 1/2$ dont le moment magnétique $\vec{\mu}$ s'exprime à l'aide des opérateurs de Pauli comme suit :

$$\vec{\mu} = g_{\mathrm{n}}\mu_{\mathrm{N}}\hbar^{-1}\vec{S} \equiv \mu\vec{\sigma} \ . \tag{21.211}$$

où $g_{\mathrm{n}} \simeq -3,82$; $\mu_{\mathrm{N}} \simeq 3,1 \times 10^{-8}$ eV/T est le magnéton de Bohr nucléaire. Le neutron est soumis à un champ magnétique constant et homogène, de module $\mathcal{B}$ définissant l'axe de quantification Oz. Dans toute la suite, on note $|\pm\rangle$ les états propres de σ_z : $\sigma_z|\pm\rangle = \pm|\pm\rangle$.

1. Quel est le Hamiltonien H ? On posera $\hbar\omega_{\mathrm{c}} = 2|\mu|\mathcal{B}$.
2. Quelles sont les énergies propres ?
3. À l'instant initial, le neutron est dans l'état :

$$|\Psi(0)\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \ . \tag{21.212}$$

   Quels sont les résultats possibles, et leurs probabilités, d'une mesure de μ_x effectuée à l'instant $t = 0+$?
4. Soit $U(t)$ l'opérateur d'évolution ; quelle est son expression en fonction de H ?
5. Écrire $U(t)$ en combinaison linéaire de l'identité $\mathbf{1}_2$ et de σ_z.
6. En déduire l'état $|\Psi(T)\rangle$ à l'instant T.
7. On mesure μ_x à l'instant T : quelles sont les probabilités $P_\pm(T)$ de trouver les valeurs $\mp\mu$?
8. On effectue une suite de mesures de μ_x aux instants $t_n = nT/N$ ($n = 1, 2, \ldots, N$). Quelle est la probabilité $P_N(T)$ pour que toutes ces mesures donnent le même résultat $+|\mu|$?
9. Quelle est la limite de $P_N(T)$ quand $N \to +\infty$? Commentez.
10. Pour un spin non-isolé (en contact avec un réservoir, un thermostat, ...), les probabilités trouvées en 7 sont affectées d'un facteur exponentiel e^{-T/T_2} représentant la relaxation transverse des composantes du moment magnétique. Dans ces conditions, que devient la probabilité $P_N(T)$ trouvée en 9 ? Conclusion.

21.6.3 À propos de la fonction de Wigner

1. En utilisant la représentation-q de l'opérateur $B(p)$, $-\mathrm{i}\hbar\delta(q-q')B(\frac{\partial}{\partial q})$, montrer directement que[103] :

$$\langle B(p)\rangle = \int_{-\infty}^{+\infty} \mathrm{d}p\, B(p) \int_{-\infty}^{+\infty} \mathrm{d}q\, W(q, p) \ . \tag{21.213}$$

[103] Pour simplifier, on pourra raisonner avec le cas pur construit sur une fonction $\Psi(q)$. Le résultat (21.213) a été obtenu en (21.175) en invoquant le théorème de convolution.

2. Soit une particule libre de masse m initialement dans le cas pur construit sur l'état gaussien :

$$\Psi(q,\,0) \stackrel{\text{déf}}{=} \frac{1}{\sqrt{\sqrt{2\pi}\Delta q}}\,\mathrm{e}^{-\frac{q^2}{4\Delta q^2}}\,\mathrm{e}^{\frac{\mathrm{i}}{\hbar}p_0 q}\ . \tag{21.214}$$

   (a) En utilisant explicitement le propagateur de la particule libre[104], montrer que la fonction de Wigner à l'instant t est donnée par :

$$W(q,\,p,\,t) = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p\xi}\,\Psi_0(q-\frac{p}{m}t-\frac{\xi}{2})\Psi_0^*(q-\frac{p}{m}t+\frac{\xi}{2})\,\mathrm{d}\xi\ . \tag{21.215}$$

   (b) En déduire l'expression de $W(q,\,p,\,t)$ et commenter.

3. Une particule de masse m et de charge Q est soumise à un champ électrique constant et uniforme de module $\mathcal{E}$, et initialement dans le cas pur construit sur l'état (21.214) avec $p_0 = 0$ (voir problème 14.6.8 du Tome I). Trouver $W(q,\,p,\,t)$, et vérifier le résultat en comparant au cas $\mathcal{E} = 0$.

4. Soit un oscillateur harmonique $(m,\,\omega)$.

   (a) L'oscillateur étant initialement dans l'état :

$$\Psi(q,\,0) = (2\pi\Delta q^2)^{-1/4}\,\mathrm{e}^{-\frac{q^2}{4\Delta q^2}}\ , \tag{21.216}$$

   utiliser l'expression de $\Psi(q,\,t)$ (voir Tome I, éq. (16.149)) pour trouver la fonction de Wigner du cas pur $|\Psi(t)\rangle\langle\Psi(t)|$.

   (b) Trouver $W(q,\,p,\,t)$ pour le cas pur formé avec l'état cohérent $|\Psi_\alpha(t)\rangle$ (voir Tome I, éq. (16.186)).

   (c) Soit $\rho(x,\,\beta\,;\,x')$ l'opérateur densité canonique ; en utilisant son expression donnée dans le Tome I, éq. (16.156), trouver la fonction de Wigner canonique $W(q,\,p,\,\beta)$. En déduire, à vue, les écarts quadratiques de la coordonnée et de l'impulsion ; examiner les limites $k_{\mathrm{B}}T \ll \hbar\omega$ et $\hbar\omega \ll k_{\mathrm{B}}T$.

21.6.4 Disparition de la cohérence spatiale pour une particule libre

Partant de la fonction de Wigner $W(q,\,p)$ relative à l'état quelconque (pur ou mixte) $\rho(q,\,q')$ (voir (21.180)), on définit sa transformée de Fourier, $C(X,\,P,\,t)$ par :

$$C(X,\,P,\,t) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^2} \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(pX+qP)}\,W(q,\,p,\,t)\,\mathrm{d}q\mathrm{d}p\ ; \tag{21.217}$$

ce que l'on peut traduire par les relations symboliques :

$$W(q,\,p,\,t) \stackrel{\mathcal{F}}{\rightarrow} C(X,\,P,\,t) \quad\Longleftrightarrow\quad C(X,\,P,\,t) \stackrel{\mathcal{F}^{-1}}{\rightarrow} W(q,\,p,\,t)\ ; \tag{21.218}$$

[104] Voir Tome I, éq. (14.129).

1. Établir les correspondances suivantes :
$$pW \overset{\mathcal{F}^{-1}}{\longrightarrow} \mathrm{i}\hbar\frac{\partial C}{\partial X} \ , \qquad \frac{\partial W}{\partial q} \overset{\mathcal{F}^{-1}}{\longrightarrow} \mathrm{i}\hbar PC \ , \qquad \frac{\partial W}{\partial p} \overset{\mathcal{F}^{-1}}{\longrightarrow} \mathrm{i}\hbar XC \ . \tag{21.219}$$

2. En déduire que, pour une particule libre de masse M, la fonction $C(X, P, t)$ satisfait l'équation suivante :
$$\frac{\partial C}{\partial t} = \frac{P}{M}\frac{\partial C}{\partial X} - \gamma X\frac{\partial C}{\partial X} - \hbar^{-2}DX^2C \ . \tag{21.220}$$

3. On définit la *cohérence globale à la distance* X [102] comme la fonction $\mathcal{C}(X)$:
$$\mathcal{C}(X, t) \overset{\text{déf}}{=} \int_{\mathbb{R}} \rho(q - \frac{X}{2}, q + \frac{X}{2}, t)\,\mathrm{d}q \ ; \tag{21.221}$$
justifier brièvement cette appellation et relier $\mathcal{C}(X, t)$ à $C(X, P, t)$

4. Soit $\psi(q)$ une fonction bien localisée de largeur typique Δq ; on définit les deux fonctions translatées $\psi_\pm \overset{\text{déf}}{=} \psi(q \mp \frac{l}{2})$, avec $\Delta q \ll l$. Discuter qualitativement l'allure de $\mathcal{C}(X, t)$ dans les deux cas suivants :
   (a) cas pur construit avec la fonction[105] $\Psi(q) \overset{\text{déf}}{=} \frac{1}{\sqrt{2}}[\psi_+(q) + \psi_-(q)]$;
   (b) cas mixte $\frac{1}{2}\sum_{\sigma=\pm 1} \psi_\sigma(q)\psi_\sigma^*(q')$.

5. Soit maintenant à résoudre l'équation pour la fonction $\mathcal{C}(X, t)$, dont l'expression initiale est désignée par $\mathcal{C}_0(X)$. Montrer que l'on peut écrire la solution sous la forme :
$$\mathcal{C}(X, t) = \mathcal{C}_0(X\mathrm{e}^{-\gamma t})\,\mathrm{e}^{-\hbar^{-2}DX^2\mu(t)} \ . \tag{21.222}$$
Écrire les équations différentielles pour les fonctions $\lambda(t)$ et $\mu(t)$, et les résoudre.

6. Il s'agit de discuter physiquement la solution qui vient d'être obtenue (pour plus de détails, voir [102]).
   (a) Écrire l'expression approchée de $\mathcal{C}(X, t)$ pour $t \ll \gamma^{-1}$, et interpréter ce résultat. En supposant que la distance l est suffisamment grande pour que l'on ait à la fois :
$$\gamma t \ll 1 \ , \qquad \frac{Dl^2}{\hbar^2}\, t \gg 1 \ . \tag{21.223}$$
   Discuter qualitativement le cas où l'état initial est le cas pur (associé à la superposition cohérente) défini en 4a ; commenter.
   (b) À l'inverse, écrire l'expression approchée de $\mathcal{C}(X, t)$ aux grands temps tels que $t \gg \gamma^{-1}$. Quelle est la valeur finale de la valeur moyenne du carré de l'impulsion $\langle p^2\rangle_{\text{eq}}$?
   (c) Supposant le bain à l'équilibre thermodynamique à la température T, trouver la relation entre la constante de diffusion D et le coefficient de frottement γ.
   (d) En déduire l'expression de $\mathcal{C}_{\text{eq}}(X) \overset{\text{déf}}{=} \mathcal{C}(X, +\infty)$, paramétrée par la longueur d'onde thermique de de Broglie $\lambda_T \overset{\text{déf}}{=} \frac{\hbar}{\sqrt{Mk_{\text{B}}T}}$.

[105] On négligera le produit scalaire $\langle\psi_+|\psi_-\rangle$.

21.6.5 Évolution de la cohérence quantique d'un atome lors de l'émission spontanée

1. *Retour sur l'opérateur densité*

   L'opérateur densité ρ fournit une description complète d'un système quantique généralisant celle que pourvoit le vecteur d'état normalisé $|\Psi\rangle$.

   (a) Dans un premier temps (cas dit *pur*), on définit ρ comme :

   $$\rho \stackrel{\text{déf}}{=} |\Psi\rangle\langle\Psi| \ . \tag{21.224}$$

   i. Combien vaut ρ^2 ? Un tel opérateur est dit *idempotent.*

   ii. À quoi est égale la trace de ρ, $\operatorname{Tr}\rho$?

   iii. Soit une observable A ; comment s'exprime à l'aide de ρ la valeur moyenne $A = \langle\Psi|A|\Psi\rangle$?

   iv. Soit $\{|a_n\rangle\}_n$ la base complète des vecteurs propres de l'observable A, sur laquelle $|\Psi\rangle$ admet la décomposition $|\Psi\rangle = \sum_n c_n|a_n\rangle$. Exprimer les éléments de matrice $\langle a_n|\rho|a_{n'}\rangle$ à l'aide des coefficients $\{c_n\}_n$.

   (b) La généralisation (cas *mixte*) consiste à définir ρ sans passer par l'intermédiaire d'un vecteur d'état, tout en maintenant pour toute valeur moyenne l'expression obtenue en 1(a)iii. Ainsi, sur la base propre d'une observable A commutant avec ρ, le développement de ρ a maintenant la forme[106] :

   $$\rho \stackrel{\text{déf}}{=} \sum_n |a_n\rangle\, p_n\, \langle a_n| \qquad (p_n \geq 0, \sum_n p_n = 1) \ . \tag{21.225}$$

   i. Quelle est la signification des p_n ? La préciser quand on raisonne avec un ensemble statistique de systèmes identiques (mais discernables).

   ii. Dans quel cas a-t-on encore $\rho^2 = \rho$?

   Dans tous les cas, les éléments diagonaux de ρ sur une base donnée sont appelés *populations*, les éléments non-diagonaux sont appelés *cohérences.*

2. *Exemple de distinction cas pur/cas mixte : deux paquets d'ondes spatialement séparés*[107]

   Il s'agit ici d'illustrer la différence fondamentale en termes de cohérences entre un cas pur et un cas mixte. Dans la suite, $\psi(x)$ désigne une fonction bien localisée autour de l'origine et ayant une largeur typique Δx ; on note $\hat{p}$ l'opérateur associé à l'impulsion.

   À partir de $\psi(x)$, on forme les deux fonctions translatées $\psi_\pm(x) \stackrel{\text{déf}}{=} \psi(x \pm a)$.

   (a) Soit $\mathcal{T}(\pm a) = \mathrm{e}^{\pm\frac{1}{\mathrm{i}\hbar}\hat{p}a}$ l'opérateur de translation de $\pm a$, et $\phi(p)$ la représentation-p de $\psi(x)$:

   $$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}px}\, \psi(x)\,\mathrm{d}x \ . \tag{21.226}$$

[106] On suppose qu'il n'y a pas de dégénérescence.

[107] Pour plus de détails, voir [102], cours II et III.

Quel est l'ordre de grandeur de la largeur Δp de $\phi(p)$?

(b) À quoi est égal $\hat{p}\phi(p)$? Et $\mathcal{T}(\pm a)\phi(p)$?

(c) En déduire que les représentations-p des états translatés $\phi_\pm(p)$ sont de la forme $\phi_\pm(p) = f_\pm(p)\,\phi(p)$ où $f_\pm(p)$ est un facteur de phase à préciser.

(d) On forme maintenant la combinaison linéaire $\Psi(x) = \frac{1}{\sqrt{2}}[\psi_+(x)+\psi_-(x)]$, dont la représentation-p est notée $\Phi(p)$. Exprimer $\Phi(p)$ à l'aide de $\phi(p)$.

(e) On suppose les deux paquets d'ondes $\psi_\pm(x)$ bien séparés dans l'espace à l'échelle a ($\Delta x \ll a$). En déduire l'allure de la distribution de probabilité de l'impulsion associée à $\Phi(p)$.

(f) Soit $\rho_{\text{pur}} = |\Psi\rangle\langle\Psi|$; préciser les valeurs typiques de x et de x' pour lesquelles les cohérences spatiales $\langle x|\rho_{\text{pur}}|x'\rangle$ sont sensiblement différentes de zéro.

(g) On considère maintenant le mélange statistique à poids égaux des deux états translatés dont la représentation-q est $\psi_\pm(x)$; dans cette représentation, on a :

$$\langle x|\rho_{\text{m},\,q}|x'\rangle = \frac{1}{2}[\psi_+(x)\psi_+^*(x') + \psi_-(x)\psi_-^*(x')] \ , \tag{21.227}$$

Que peut-on dire des cohérences spatiales $\langle x|\rho_{\text{m},\,q}|x'\rangle$?

(h) En représentation-p, le même mélange est décrit par :

$$\langle p|\rho_{\text{m},\,p}|p'\rangle = \frac{1}{2}[\phi_+(p)\phi_+^*(p') + \phi_-(p)\phi_-^*(p')] \ , \tag{21.228}$$

dont les éléments diagonaux en représentation-p $\langle p|\rho_{\text{m},\,p}|p\rangle$ donnent la densité de probabilité de l'impulsion. Tracer le graphe de $\langle p|\rho_{\text{m},\,p}|p\rangle$. Comparer à celui obtenu en 2e pour l'état pur alors considéré.

3. *Ce que fait une mesure d'observable*

On effectue une mesure de l'observable A sur *un* système dans l'état $|\Psi\rangle$, et on trouve la valeur a_{n_0} ; selon l'interprétation orthodoxe, la réduction du paquet d'ondes se traduit par l'évolution non unitaire $|\Psi\rangle \longrightarrow |a_{n_0}\rangle$.

(a) Soit un ensemble statistique de systèmes identiques discernables tous décrits par le même vecteur $|\Psi\rangle$; on effectue simultanément la mesure de A sur chacun des systèmes de l'ensemble. Reprenant les mêmes notations qu'en 1(a)iv, préciser :

i. l'opérateur densité juste avant la mesure, $\rho_{\text{juste avant}}$.

ii. l'opérateur densité juste après la mesure, $\rho_{\text{juste après}}$.

iii. la valeur moyenne des valeurs obtenues pour l'observable A.

(b) D'une phrase, résumer l'effet d'une mesure idéale sur les cohérences quantiques.

4. *Opérateur densité réduit*

   Pour un système donné, considérons deux sortes de degrés de liberté, notés respectivement $\{a\}$ et $\{\alpha\}$, associés à deux catégories d'observables ayant les vecteurs propres $|a_n\rangle \otimes |\alpha_\nu\rangle \equiv |a_n,\, \alpha_\nu\rangle$.

   Très souvent, on ne s'intéresse qu'à une partie des degrés de liberté d'un système, ceux ici notés $\{a\}$. La trace partielle sur les autres degrés de liberté $\{\alpha\}$ permet de définir un opérateur densité *réduit* $\tilde{\rho}$, défini comme :

$$\tilde{\rho} \stackrel{\text{déf}}{=} \sum_\nu |a_n,\, \alpha_\nu\rangle \, \langle a_n,\, \alpha_\nu|\rho|a_{n'},\, \alpha_\nu\rangle\langle a_{n'},\, \alpha_\nu| \ . \tag{21.229}$$

   Même si l'opérateur relatif à tous les degrés de liberté est idempotent, en général l'opérateur réduit ne l'est pas, comme le montre l'exemple suivant.

   (a) Soit deux particules *discernables* de spin 1/2, notées 1 et 2 si nécessaire. On désigne par $|\uparrow\rangle$ et $|\downarrow\rangle$ les états propres de $\vec{S}^{\,2}$ et S_z de chaque spin.
   Donner l'expression de l'état singulet, et celle de l'opérateur densité ρ associé.

   (b) On mesure le spin de la particule 1, et on trouve $+\hbar/2$. Que peut-on en déduire sur le spin de la particule 2 ? Qu'en pensez-vous ?

   (c) Déduire de 4a la matrice de l'opérateur densité réduit $\tilde{\rho}$ sur la base où ne figurent plus que les degrés de liberté de la particule 1 (le donner par sa matrice sur la base $(|\uparrow\rangle,\, |\downarrow\rangle)$. Est-ce un cas pur ?

   (d) Quelle est la signification physique précise des éléments diagonaux de $\tilde{\rho}$?

5. *Évolution d'un état discret couplé à un continuum*

   Pour un système isolé de Hamiltonien H, l'opérateur d'évolution est $U(t) = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Ht}$. Si $\{|E_n\rangle\}_n$ est l'ensemble des vecteurs propres de H, $U(t)$ s'écrit :

$$U(t) = \sum_n |E_n\rangle \, \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}E_n t}\langle E_n| \ . \tag{21.230}$$

   (a) On définit la *résolvante* $G(z)$ par la relation :

$$G(z) \stackrel{\text{déf}}{=} \frac{1}{z\mathbf{1} - H} \ . \tag{21.231}$$

   Soit l'opérateur $U_+(t)$:

$$U_+(t) = \frac{1}{2\mathrm{i}\pi}\int_{C_+} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}zt}G(z)\,\mathrm{d}z \ , \tag{21.232}$$

   où C_+ est la droite située juste au-dessus de l'axe réel parcourue de $+\infty + \mathrm{i}\varepsilon$ à $-\infty + \mathrm{i}\varepsilon$, où $\varepsilon \in \mathbb{R}_+$. À l'aide du théorème des résidus[108], montrer que $U_+(t) = U(t)$ si $t > 0$, $U_+(t) = 0$ si $t < 0$.

[108] On supposera la convergence à l'infini assez rapide pour que le lemme de Jordan soit applicable.

(b) Dans le cas où H se décompose en $H_0 + V$, on peut aussi définir :

$$G_0(z) \stackrel{\text{déf}}{=} \frac{1}{z\mathbf{1} - H_0} \; ; \tag{21.233}$$

en prenant garde à la non-commutation de H_0 et V, établir l'équation de Dyson $G(z) = G_0(z) + G_0(z)VG(z)$.

(c) H_0 est supposé avoir des états propres de deux sortes :

i. un état $|\mathrm{r}\rangle$, d'énergie $E_\mathrm{r} > 0$;

ii. des états $|E\rangle$ dont l'énergie E varie continûment de 0 à $+\infty$.

Sur cette base, V est purement non-diagonal ($\langle E|V|E'\rangle = \langle \mathrm{r}|V|\mathrm{r}\rangle = 0$) ; posant $\langle E|V|\mathrm{r}\rangle = v(E)$, itérer une fois l'équation de Dyson pour montrer que l'élément de matrice $\langle \mathrm{r}|G(z)|\mathrm{r}\rangle$ peut se mettre sous la forme :

$$\langle \mathrm{r}|G(z)|\mathrm{r}\rangle = \frac{1}{z - E_\mathrm{r} - \Sigma(z)} \stackrel{\text{déf}}{=} \mathcal{G}(z) \; . \tag{21.234}$$

où la *self-energy* $\Sigma(z)$ a pour expression :

$$\Sigma(z) = \int_0^{+\infty} \frac{|v(E)|^2}{z - E}\,\mathrm{d}E \; . \tag{21.235}$$

(d) Le système étant préparé à $t = 0$ dans l'état $|\mathrm{r}\rangle$, soit $|\Psi(t)\rangle$ l'état à l'instant $t > 0$; posant $z = E + \mathrm{i}\varepsilon$ ($E \in \mathbb{R}$, $\varepsilon > 0$) et $\Sigma_+(E) \stackrel{\text{déf}}{=} \lim_{\varepsilon\to 0}\Sigma(E + \mathrm{i}\varepsilon)$, écrire l'amplitude $A(t) \stackrel{\text{déf}}{=} \langle \mathrm{r}|\Psi(t)\rangle$ sous la forme d'une intégrale réelle.

(e) Dans l'hypothèse d'un couplage faible ("V petit"), on remplace désormais $\mathcal{G}(E + \mathrm{i}0)$ par la résolvante approchée $\mathcal{G}_\mathrm{ap}(E)$:

$$\mathcal{G}(E + \mathrm{i}0) \longrightarrow \mathcal{G}_\mathrm{ap}(E) = \frac{1}{E - E_\mathrm{r} - \Sigma_+(E_\mathrm{r})} \; . \tag{21.236}$$

Exprimer $\Sigma_+(E_\mathrm{r})$ en fonction de[109] :

$$\Delta \stackrel{\text{déf}}{=} \hbar^{-1}\mathcal{P}\int_0^{+\infty} \frac{|v(E)|^2}{E_\mathrm{r} - E}\,\mathrm{d}E \; , \qquad \Gamma \stackrel{\text{déf}}{=} \frac{2\pi}{\hbar}|v(E_\mathrm{r})|^2 \; . \tag{21.237}$$

(f) Des résultats précédents, déduire l'expression (approchée) de l'amplitude $A(t)$.

(g) L'état $|\Psi(t)\rangle$ peut ainsi s'écrire :

$$|\Psi(t)\rangle = A(t)|\mathrm{r}\rangle + \int \mathrm{d}E\, c_E(t)\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Et}\,|E\rangle \; . \tag{21.238}$$

Écrire l'équation différentielle satisfaite par les coefficients $c_E(t)$, et en donner la solution.

(h) En déduire l'expression approchée de $|\Psi(t)\rangle$ pour $t \gg \Gamma^{-1}$.

[109] On rappelle que $\lim_{\varepsilon\to 0_+} \frac{1}{x+\mathrm{i}\varepsilon} = \mathcal{P}\frac{1}{x} - \mathrm{i}\pi\delta(x)$; $\mathcal{P}$ désigne la partie principale de Cauchy.

6. *Évolution de la cohérence d'un atome lors de l'émission spontanée*

Dans cette section, on considère un atome de masse M à deux états internes : $|\mathrm{g}\rangle$ (fondamental, énergie prise comme zéro des énergies) et $|\mathrm{e}\rangle$ (état excité, d'énergie $E_\mathrm{e} \equiv \hbar\omega_\mathrm{e}$). On désigne par $|\vec{K}\rangle$ les états propres de $\vec{P}$, impulsion du centre de masse de l'atome.

Cet atome est couplé au champ de rayonnement dont on n'introduit que deux types d'états : l'état fondamental $|\mathrm{vide}\rangle$ (zéro photon) et les états[110] $|\vec{k}\rangle$ à un photon d'impulsion $\vec{k}$ et d'énergie $\hbar\omega$. Cela fait, on introduit la base d'états :

$$|\vec{K}\rangle \stackrel{\text{déf}}{=} |\mathrm{e}\rangle \otimes |\vec{K}\rangle \otimes |\mathrm{vide}\rangle \ , \qquad |\vec{k},\, \vec{K}\,'\rangle \stackrel{\text{déf}}{=} |\mathrm{g}\rangle \otimes |\vec{K}\,'\rangle \otimes |\vec{k}\rangle \ . \tag{21.239}$$

Soit $H \stackrel{\text{déf}}{=} H_0 + V$ le Hamiltonien du système atome + champ ; on note :

$$w(\vec{k},\, \vec{K}) \stackrel{\text{déf}}{=} \langle \vec{k},\, \vec{K} - \vec{k}|V|\vec{K}\rangle \ . \tag{21.240}$$

(a) Quelles sont les expressions des énergies (à l'ordre zéro en V) des états $|\vec{K}\rangle$ et $|\vec{k},\, \vec{K}\rangle$?

(b) Quelle est la relation entre $\vec{k}$, $\vec{K}$ et $\vec{K}\,'$ pour deux états $|\vec{K}\rangle$ et $|\vec{k},\, \vec{K}\,'\rangle$ effectivement couplés par V ?

(c) Par référence avec la partie 5, préciser les états qui vont ici jouer les rôles respectifs de $|\mathrm{r}\rangle$ et $|E\rangle$.

(d) On introduit à nouveau des quantités Γ et Δ comme en (21.237), et on suppose qu'à $t = 0$, le système est dans l'état $|\vec{K}_0\rangle$. En utilisant le résultat obtenu en 5h, montrer que pour $t \gg \Gamma^{-1}$ l'expression approchée du vecteur d'état $|\Psi(t)$ est de la forme :

$$|\Psi(t)\rangle \simeq \sum_{\vec{k}} \gamma(\vec{k},\, \vec{K}_0)\, \mathrm{e}^{-\mathrm{i}[\omega + \frac{\hbar(\vec{k}-\vec{K}_0)^2}{2M}]\, t}\, |\vec{K}_0 - \vec{k},\, \vec{k}\rangle \ , \tag{21.241}$$

avec :

$$\gamma(\vec{k},\, \vec{K}_0) = \frac{w(\vec{k},\, \vec{K}_0)}{\hbar(\omega - \omega_0 + \frac{\mathrm{i}}{2}\Gamma) - \frac{\hbar^2}{M}\,\vec{k}.\vec{K}_0} \ , \tag{21.242}$$

où ω_0 est une pulsation à préciser. Comment s'appelle l'effet associé au terme $\vec{k}.\vec{K}_0$?

(e) Quand on ne s'intéresse qu'au mouvement du centre de masse de l'atome (mouvement *externe*), il suffit de connaître l'opérateur densité réduit correspondant, noté ρ_ext. En normalisant toute onde plane suivant $\langle \vec{r}\,|\vec{K}\rangle \stackrel{\text{déf}}{=} L^{-3/2}\, \mathrm{e}^{\mathrm{i}\vec{K}.\vec{r}}$, calculer la cohérence spatiale $\langle \vec{r}\,|\rho_\mathrm{ext}(t = 0)|\vec{r}\,'\rangle$ avant émission d'un photon. Commenter.

(f) En raisonnant dans le repère du centre de masse de l'atome avant émission, écrire l'expression formelle de l'opérateur densité réduit $\rho_\mathrm{ext}(t)$ à $t \gg \Gamma^{-1}$, après émission d'un photon $\hbar\vec{k}$. Est-ce un cas pur ou un cas mixte ?

[110] Pour simplifier, on délaisse la polarisation du photon.

(g) Déduire du résultat précédent l'expression de $\langle\vec{r}\,|\rho_{\text{ext}}(t\gg\Gamma^{-1})|\vec{r}\,'\rangle$ en fonction des coefficients $\gamma(\vec{k},\,\vec{0})$.

(h) Il s'agit maintenant de calculer effectivement la cohérence $\langle\vec{r}\,|\rho_{\text{ext}}(t)|\vec{r}\,'\rangle$ pour $t\gg\Gamma^{-1}$. En raison de la variation résonnante de γ pour $\omega\sim\omega_0$, on ignore désormais toute variation en $\vec{k}$ de l'élément de matrice $w(\vec{k},\,\vec{0})$, notant simplement w la valeur constante retenue.

i. Montrer que, après intégration sur les angles, $\langle\vec{r}\,|\rho_{\text{ext}}|\vec{r}\,'\rangle$ est proportionnelle à l'intégrale I :

$$I \stackrel{\text{déf}}{=} \int_0^{+\infty}\frac{k\sin kl}{(k-k_0)^2c^2+\frac{\Gamma^2}{4}}\,\mathrm{d}k \qquad (l\stackrel{\text{déf}}{=}\|\vec{r}-\vec{r}\,'\|\ ,\ k_0=\frac{\omega_0}{c})\ . \quad (21.243)$$

ii. En estimant l'ordre de grandeur de $\frac{\Gamma}{c}$ pour une transition optique typique, justifier rapidement que l'on rejette la borne inférieure de l'intégrale à $-\infty$; en déduire l'expression de $\langle\vec{r}\,|\rho_{\text{ext}}|\vec{r}\,'\rangle\equiv\langle\vec{r}\,|\rho_{\text{ext}}|\vec{r}\,\rangle\,C(l)$.

iii. Tracer l'allure de la variation de $C(l)$ en fonction de l, et commenter.

iv. $C(l)$ est essentiellement le produit de deux facteurs :

- le premier est la transformée de Fourier d'une fonction-porte. Expliquer pourquoi on peut le comprendre comme une conséquence du caractère aléatoire de la direction d'émission du photon ;
- le second est une exponentielle. Estimer l'ordre de grandeur de la longueur caractéristique d'extinction.

(i) On en revient à l'état dont l'expression est donnée en (21.241), avec toujours $\vec{K}_0=\vec{0}$. En utilisant la relation de fermeture $\sum_{\vec{r}}|\vec{r}\,\rangle\langle\vec{r}\,|$, montrer que l'état du système atome + champ se met sous la forme :

$$|\Psi(t)\rangle = \sum_{\vec{r}}|\vec{r}\,\rangle\otimes|\chi_{\vec{r}}(t)\rangle\ , \qquad (21.244)$$

et donner l'expression de $|\chi_{\vec{r}}(t)\rangle$.

(j) Expliquer en quoi il existe, après émission, une corrélation entre la position de l'atome et l'état du champ, de même nature que celle obtenue dans le point 4b.

Quel est l'état du champ issu d'une mesure de la position de l'atome à l'instant t ayant trouvé l'atome au point $\vec{r}_0$?

(k) Inversement, on peut considérer l'état du champ $|\chi_{\vec{r}}\rangle$ comme un *pointeur* ("l'aiguille de l'ampèremètre") donnant la position de l'atome. Si la résolution de ce pointeur était parfaite, on aurait $\langle\chi_{\vec{r}\,'}|\chi_{\vec{r}\,''}\rangle=0$ si $\vec{r}\,'\neq\vec{r}\,''$ puisque les deux états $|\vec{r}\,'\rangle$ et $\vec{r}\,''\rangle$ sont orthogonaux.

Qu'en est-il ? Montrer que le champ "mesure" la position de l'atome avec une résolution de l'ordre de la longueur d'onde $\lambda_0=\frac{c}{2\pi\omega_0}$.

On rencontrera l'intégrale $\int_{-\infty}^{+\infty}\frac{k\,\mathrm{e}^{\mathrm{i}kl}}{(k-k_0)^2+\alpha^2}\,\mathrm{d}k$, qui se calcule aisément par résidus.

physique

Chapitre 22

Particules identiques

Ce chapitre expose les problèmes particuliers que soulève en Mécanique quantique la description des systèmes composés de particules identiques, débouchant sur la nécessité d'un nouveau postulat. Après l'énoncé de celui-ci, on montrera le rôle joué par les opérateurs de permutation, avant une initiation à la Seconde Quantification.

22.1 Indiscernabilité des particules identiques en Mécanique quantique

En se donnant comme objectif la description des grandeurs observables, conformément aux prolégomènes de Heisenberg [164], la Mécanique quantique ne considère pas certains concepts classiques, sans toutefois exclure de les voir réapparaître dans la bonne limite. Il en va ainsi de la notion de trajectoire, dont on a vu qu'elle est impossible à maintenir ne serait-ce qu'en vertu des inégalités de Heisenberg, mais qui ressort en bout de course, quoique de façon très singulière, dans la formulation de Feynman quand on y fait tendre la constante de Planck vers zéro (voir Tome I, chapitre 14, section 14.3, notamment p. 502).

L'absence, ou en tout cas l'inexistence dans le formalisme, de la notion de trajectoire, fait surgir un problème inédit quand il est question de particules *identiques*. En Mécanique Classique, la description d'un système composé de particules identiques ne pose aucun problème particulier, puisqu'il est possible de numéroter (ou de colorier) conventionnellement les particules dans l'état initial, puis de les suivre une à une sur leur trajectoire. À tout instant, on peut ainsi identifier par son dossard la particule portant

au départ un numéro donné ; c'est bien ce suivi de proche en proche qui fait de leur numéro (ou de leur couleur) une "constante du mouvement" individuel.

En Mécanique quantique, il en va tout autrement : comme il n'existe plus de trajectoire, on ne peut plus suivre les particules à la trace. Même s'il est encore parfois possible de les numéroter au départ, à condition qu'elles se trouvent dans des régions distinctes de l'espace, aucun procédé n'existe pour retrouver leur numéro dans l'état final si leurs paquets d'ondes se sont enchevêtrés entre-temps[1], c'est-à-dire après qu'elles ont, d'une façon ou d'une autre, interagi les unes avec les autres. L'impossibilité de distinguer deux particules identiques disparaît dans la limite classique : alors, les paquets d'ondes restent très bien localisés et ne se mélangent pas, leur centre suit en gros la trajectoire classique, qui reprend alors une grande part de son sens physique – le non-recouvrement entre les paquets d'ondes restitue *de facto* la discernabilité classique.

De cette *indiscernabilité* résulte notamment qu'une question comme *"Quelle est la probabilité de trouver l'électron 1 en* $\vec{r}_1$ *et l'électron 2 en* $\vec{r}_2$ *?"* n'a pas de signification ; au contraire, la question *"quelle est la probabilité de trouver un électron en* $\vec{r}_1$ *et un en* $\vec{r}_2$*"* est sensée. L'impossibilité de distinguer plusieurs particules identiques cohabitant au sein d'un système (ou ayant cohabité) – d'où l'invariance des propriétés physiques dans tout *échange* des variables dynamiques des particules – se révèle comme toujours par une certaine *symétrie*, ici il s'agit d'une symétrie de *permutation* de l'objet décrivant l'état du système, sa fonction d'onde par exemple. Cette invariance est traduite par un nouveau postulat aux conséquences extraordinaires, toutes reposant sur le fait qu'il conduit automatiquement à la définition de deux grandes classes de particules, les *fermions* (spin demi-entier) – pour lesquels fut énoncé le *Principe de Pauli* – et les *bosons* (spin entier). Chaque famille obéit à une statistique radicalement différente de celle à laquelle l'autre est soumise (Fermi - Dirac vs. Bose - Einstein), donnant lieu à des comportements collectifs aux différences spectaculaires ; très schématiquement : les fermions n'aiment pas cohabiter dans la même case quantique, les bosons au contraire ont tendance à se rassembler (d'où, à basse température, le phénomène de condensation de Bose).

Ce postulat est, ni plus, ni moins, l'un des fondements de l'explication de la stabilité de la matière constituée. C'est parce que la matière est formée de fermions – les bosons étant des particules associées aux interactions (champs) – que la limite thermodynamique existe ([24], § 2.2.1) ; en particulier, pour qu'un système de particules de charges opposées soit stable, il faut que l'une des espèces soit constituée de fermions. Le Principe de Pauli sauve la situation : sans lui, l'énergie de l'état fondamental varierait comme $N^{7/5}$ en fonction du nombre N de particules et la limite thermodynamique (qui exige que l'énergie soit extensive, $E \propto N$) tout simplement n'existerait pas. S'il s'agit là de l'une de ses conséquences les plus spectaculaires, d'autres ne sont pas moins importantes. Ce

[1] En revanche, si à tout instant les paquets d'ondes restent bien séparés et ne se mélangent pas, la trajectoire classique reprend alors son sens : s'ils ne s'étalent pas trop et si le potentiel ne varie pas trop vite dans l'espace, leur centre suit en gros la trajectoire classique (Théorème d'Ehrenfest). En définitive, il existe heureusement une sorte de continuité vers la situation classique : le non-recouvrement entre les paquets d'ondes rétablit la discernabilité classique et l'impossibilité de distinguer deux particules identiques disparaît – voir à ce sujet la discussion de Messiah [15] *"Faut-il toujours antisymétriser la fonction d'onde ?"*

postulat permet aussi de rendre compte des sources du magnétisme (une énigme de plus résolue par la théorie quantique – voir le théorème de Miss van Leeuwen (Tome I, chapitre 11, note 1) : même en l'absence de toute interaction (au sens classique) de nature magnétique, il entraîne que l'énergie d'un atome dépend de la valeur de son spin total (voir chapitre 27, en particulier section 27.7).

La motivation première de Pauli[2] pour énoncer le *Principe d'exclusion de Pauli* (1923) était de rationaliser le tableau de Mendeleiev, sur des bases alors purement empiriques. Cette affirmation de principe énoncée à l'aube de la Mécanique quantique se trouve être en réalité une conséquence du fait que les électrons au sein d'un même atome sont des fermions indiscernables. La structure des atomes, et donc la classification périodique, est ainsi en réalité une autre conséquence de ce postulat.

Il en va de même de la distinction entre isolants et conducteurs, résultat de l'identité des électrons dans un métal et de leur sujétion à la statistique de Fermi - Dirac. Il convient d'ajouter que la nature fermionique ou bosonique d'un système composite dépend très subtilement des interactions entre ses composants élémentaires : à (très) basse température (usuellement $T \lesssim 5$ K), les électrons d'un métal peuvent s'associer par paires[3] ayant toutes les caractéristiques de bosons[4] (*paires de Cooper*). Ces bosons *non-élémentaires*, immergés dans un fluide d'électrons "normaux", peuvent en un sens se condenser pour produire l'état supraconducteur (voir chapitre 32, section 32.5). La condensation de Bose, sous sa forme historique très vite reconnue ou par ses manifestations récentes avec des atomes froids, est un autre phénomène spectaculaire que seule l'indiscernabilité (et ce qu'elle exige) permet de comprendre. Aussi différents soient-ils, tous ces phénomènes sont liés au fait que deux particules identiques sont, *de facto*, indiscernables l'une de l'autre.

---

[2]Wolfgang Ernst Pauli (1900-1958) était un physicien autrichien, dont le deuxième prénom fut choisi en hommage à son grand-père maternel, Ernst Mach (1838-1916), auteur d'un principe dont Feynman dit avec malice [165] : *"Pour autant que nous le sachions, Mach a raison : personne n'a à ce jour démontré l'inexactitude de son principe en supprimant tout l'univers pour constater ensuite qu'une masse continuait éventuellement à avoir une inertie !"*.

Les contributions théoriques magistrales de Pauli lui ont valu le Nobel en 1945 avec la mention *"for the discovery of the Exclusion Principle, also called the Pauli Principle"*. Réputé pour être particulièrement critique et exigeant, voire arrogant (ce dont Ehrenfest fut victime), on lui doit ce commentaire *"it's not even wrong"* à propos d'un article qu'il jugeait sévèrement. L'*effet Pauli* désigne un prétendu fait empirique selon lequel la présence de Pauli dans un laboratoire expérimental faisait échouer les manips'..., d'où l'*autre* Principe d'exclusion de Pauli, énoncé par Stern (voir Tome I, note 16 p. 373)...

[3]On peut trouver énigmatique que deux électrons, qui suivant la loi de Coulomb se repoussent fortement, peuvent néanmoins s'associer pour former un édifice stable. Suivant la théorie standard, dite théorie BCS, il existe de surcroît une interaction *attractive* colportée par les vibrations du réseau (*phonons*, voir chapitre 31) permettant la formation de paires liées, dont on peut se faire la représentation physique suivante. Lorsqu'un électron se propage dans le réseau d'ions, il le déforme ; la déformation du réseau entraînant avec elle le fond continu de densité électronique, une lacune positive apparaît, *via* la conservation locale de la charge. Un autre électron survenant alors est attiré par cette polarisation positive et le résultat net est bien une interaction attractive dynamique entre les deux électrons, celui qui a créé la lacune positive et celui qui est attirée par elle ; le fluide normal où baignent ces deux électrons joue un rôle crucial : notamment, c'est lui qui *écrante* leur répulsion nue. Si la supraconductivité n'existe qu'à basse température, c'est parce que les paires de Cooper sont fragiles et se cassent dès que l'énergie thermique $k_{\mathrm{B}}T$ devient comparable à leur énergie de liaison (voir chapitre 32, section 32.5).

[4]Il faut évidemment un nombre *pair* de fermions pour former un spin entier.

Deux particules sont dites *identiques* si toutes leurs grandeurs intrinsèques (masse, charge, spin, etc.) sont les mêmes. Un proton et un électron ont même spin mais sont distincts par leurs masses. Un électron et un positron ont même masse et même spin mais ont des charges opposées. Aucune expérience ne permet de distinguer deux particules identiques, sauf si l'on sait qu'elles sont – et restent à demeure – dans deux régions de l'espace dont l'intersection est vide. Cette *identité* entraîne automatiquement qu'aucune propriété d'un système contenant des particules identiques n'est modifiée si l'on procède à l'échange des variables de deux d'entre elles.

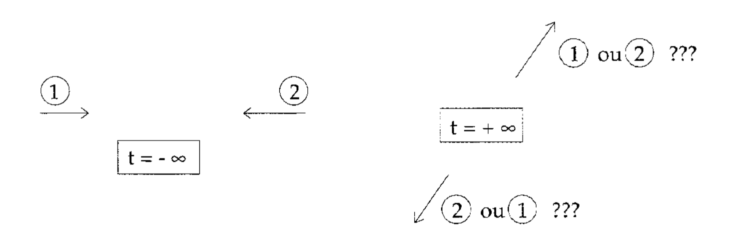


Figure 22.1: Représentation schématique de la collision de deux particules *identiques*.

Afin de situer la discussion sur un plan concret, considérons la collision, observée dans le repère du centre de masse, de deux particules identiques venant à la rencontre l'une de l'autre (voir fig. 22.1). Lorsqu'elles sont à l'infini l'une de l'autre, il est loisible de numéroter *conventionnellement* 1 celle qui vient de gauche et 2 celle qui vient de droite. Après la collision, l'observation de l'une d'entre elles porte sur son énergie, son impulsion, son angle de diffusion, etc. Mais l'identification s'arrête ici : on ne peut pas retrouver son numéro puisqu'il est impossible de retracer sa trajectoire et d'en déduire que la particule observée est celle venue d'un côté ou de l'autre. Tant que les deux paquets d'ondes sont distincts, on peut schématiser les déplacements de leurs centres ; dès qu'ils s'interpénètrent, il n'y a plus deux paquets d'ondes identifiables mais une zone diffuse dans l'espace où la probabilité de trouver les *deux* particules est non-nulle. Après séparation, le "flou" lié à l'imbrication des paquets d'ondes (une forme d'*intrication*) lors de la collision[5] introduit une irréversibilité – en ce qui concerne l'identification des particules –, et rend impossible toute autre identification que les grandeurs physiques intrinsèques de chaque particule (masse, charge, spin, etc.), lesquelles sont identiques puisque précisément les particules le sont. L'état final est certes parfaitement déterminé – le système évolue par l'équation de Schrödinger et l'état initial peut être défini sans ambiguïté –, mais aucun moyen n'existe plus de reconnaître la *couleur* conventionnelle de chacune des deux particules.

Un autre exemple ([25], § XIV. 3) permet de saisir comment l'impossibilité de distinguer deux particules identiques conduit à une difficulté majeure si l'on s'en tient aux six postulats déjà énoncés (voir Tome I, chapitre 12). Soit deux particules identiques de

[5]La situation ici décrite ressemble à celle conduisant au fameux paradoxe EPR. En fait, celui-ci est d'une tout autre nature, et ne dépend nullement du fait que les particules sont identiques ou non. *A priori*, la notion de *non-séparabilité* est indépendante de la notion d'*indiscernabilité*.

spin $S = \frac{1}{2}$, dont on ignore pour simplifier les degrés de liberté orbitaux. L'espace des états est de dimension 4, et est engendré par les vecteurs $|m_1, m_2\rangle$ avec $m_i = \pm\frac{1}{2}$. En l'absence de tout autre degré de liberté, le spin de chaque particule est la seule et unique observable individuelle pouvant faire l'objet d'une mesure. En ce sens, connaître les deux spins constitue l'information maximale permise par la Mécanique quantique, telle qu'elle a été exposée jusqu'à maintenant.

Supposons que l'on vient de faire une mesure individuelle de S_{1z} et de S_{2z} qui a donné $+\frac{\hbar}{2}$ pour un spin, $-\frac{\hbar}{2}$ pour l'autre[6]. On peut donc affirmer que l'un des spins est dans l'état $+\frac{1}{2}$, l'autre dans l'état $-\frac{1}{2}$ et, juste après cette mesure ayant réduit le paquet d'ondes, l'état devrait en principe être complètement déterminé. Or celui-ci peut être $|+\frac{1}{2}, -\frac{1}{2}\rangle$, ou $|-\frac{1}{2}, +\frac{1}{2}\rangle$, ou n'importe quelle combinaison linéaire des deux. Il est en effet certain que tout état du genre :

$$|\chi\rangle = a|+\frac{1}{2}, -\frac{1}{2}\rangle + b|-\frac{1}{2}, +\frac{1}{2}\rangle \ , \qquad |a|^2 + |b|^2 = 1 \ . \tag{22.1}$$

sur lequel on refait immédiatement une mesure de S_{1z} et de S_{2z} donne avec probabilité 1 ce qu'a donné la première mesure – et l'on ne doit rien exiger de plus.

Les postulats énoncés jusqu'à présent ne permettent donc pas de fixer[7] les coefficients a et b, de sorte que tout état du type (22.1) à la fois représente la connaissance maximale du système et est dégénéré (la dégénérescence correspondante est appelée dégénérescence d'échange) – on peut même dire qu'il est *totalement* indéterminé. Une objection à cet argument consiste à dire qu'il existe en fait un moyen de fixer a et b : il suffit de procéder à une mesure du spin *total* $\vec{S} = \vec{S}_1 + \vec{S}_2$. Si l'on a trouvé $S = 1$ alors l'état issu de la mesure est l'état triplet pour lequel $a = b = 1/\sqrt{2}$ (on additionne les deux moments cinétiques et on détermine les composantes de l'état triplet par application de S_- à $|+\frac{1}{2}, +\frac{1}{2}\rangle$) ; si on a trouvé $S = 0$, c'est l'état singulet, caractérisé par $a = -b = 1/\sqrt{2}$. Ce faisant, la connaissance acquise porte sur un système global de deux particules, où les caractéristiques individuelles de chacune se fondent dans le collectif. De surcroît, si l'on se demande quelle est maintenant la probabilité de trouver deux spins opposés pour les deux projections le long de Oz, sur quel état doit-on projeter l'état issu de la mesure du spin total ? $|+\frac{1}{2}, -\frac{1}{2}\rangle$ ou $|-\frac{1}{2}, +\frac{1}{2}\rangle$, ou une (quelle ?) combinaison linéaire ?

L'impossibilité de savoir ce que sont les coefficients a et b, et donc l'ambiguïté sur l'état initial, interdit évidemment de faire la moindre prévision sur le système. Par exemple, soit à trouver la probabilité $P(a, b)$ pour qu'une mesure simultanée de S_{1x} et S_{2x}, effectuée juste après la préparation (!?) fabriquant l'état (22.1), donne la valeur $+\frac{\hbar}{2}$ pour chaque spin. L'état propre $|+\frac{1}{2}\rangle_x$ de S_x se décompose comme suit sur les états propres de S_z, $|\pm\frac{1}{2}\rangle$:

$$|\frac{1}{2}\rangle_x = \frac{1}{\sqrt{2}}\left(|+\frac{1}{2}\rangle + |-\frac{1}{2}\rangle\right) \ . \tag{22.2}$$

[6] Les deux opérateurs S_{1z} et S_{2z} commutent puisqu'ils agissent sur des variables différentes.

[7] L'arbitraire vient précisément de l'identité des particules : toute action extérieure (au moyen d'un champ par exemple) ne peut produire un effet sélectif puisque tous les attributs intrinsèques des deux particules sont strictement les mêmes.

Pour les deux spins, le produit tensoriel se développe suivant :

$$|\frac{1}{2}\rangle_{1x} \otimes |\frac{1}{2}\rangle_{2x} = \frac{1}{2}\Big(|+\frac{1}{2}, +\frac{1}{2}\rangle + |+\frac{1}{2}, -\frac{1}{2}\rangle + |-\frac{1}{2}, +\frac{1}{2}\rangle + |-\frac{1}{2}, -\frac{1}{2}\rangle\Big) \ . \qquad (22.3)$$

Suivant les postulats déjà énoncés, la probabilité $P(a, b)$ de trouver $+\frac{\hbar}{2}$ pour chaque spin le long de Ox en effectuant une mesure sur l'état $|\psi\rangle$ (22.1) est le module au carré du produit scalaire :

$$P(a, b) = \big|(|\chi\rangle, |\frac{1}{2}\rangle_{1x} \otimes |\frac{1}{2}\rangle_{2x})\big|^2 \ . \qquad (22.4)$$

Le calcul à partir de (22.3) et (22.1) donne immédiatement :

$$\boxed{P(a, b) = \big|\frac{1}{2}(a+b)\big|^2} \qquad (22.5)$$

Sans surprise, cette probabilité est totalement *indéterminée* puisqu'elle dépend des deux coefficients arbitraires a et b : en raison de l'identité des particules, et en l'état actuel des choses, aucun moyen n'existe pour fixer ces derniers afin de fixer l'état de départ : cette ambiguïté conduit à l'impossibilité de prévoir le résultat de la deuxième mesure, comme de faire toute autre prévision d'ailleurs.

Cette difficulté peut être résolue en adoptant une prescription, qui sera en fait un postulat supplémentaire devant permettre de fixer a et b et d'obtenir alors une réponse unique à la question posée. L'énoncé de ce postulat est donné au paragraphe suivant ; avant d'en venir là, la remarque suivante est utile.

Soit un système de N particules identiques ; oubliant leurs possibles interactions mutuelles, l'état de chacune d'entre elles est complètement défini par un jeu de nombres quantiques, α. L'état du système sera spécifié complètement en disant qu'il y a n_1 particules dans l'état α_1, n_2 dans l'état α_2, etc. On ne peut pas en dire plus en précisant par exemple que les n_1 premières sont dans α_1, les n_2 suivantes dans α_2 etc. Pour un ensemble d'électrons, il est possible de calculer la probabilité de trouver *un* électron en un certain point avec un spin donné, mais c'est n'importe lequel des électrons constituant le système ; l'identification ne peut être poussée plus loin. En définitive, pour des particules identiques sans interaction, les bonnes variables quantiques sont les *nombres d'occupation* des différents états à une particule[8]. Cette reformulation porte le nom (impropre) de *Seconde quantification*, dont une présentation élémentaire sera donnée plus loin (section 22.5). Indépendamment du fait qu'elle prend naturellement en compte l'indiscernabilité des particules identiques, elle permet de décrire les systèmes dont le nombre de particules n'est pas fixé – un impératif pour la Théorie quantique des champs –, et se révèle de surcroît d'une grande commodité technique.

[8]Ce formalisme n'est heureusement pas caduc en présence d'interactions. Si celles-ci sont faibles, on parle alors de quasi-particules, ou de particules habillées. Si elles sont fortes, on redéfinit de nouveaux modes propres pour lesquels il existe à nouveau des nombres d'occupation.

22.2 Le postulat de symétrisation

La difficulté stigmatisée dans la section précédente est éliminée par l'adoption du postulat suivant :

Postulat 7 *Pour un système contenant N particules identiques, les seuls états physiques sont soit symétriques (pairs) soit antisymétriques (impairs) dans l'échange de deux quelconques de ces particules.*

Autrement dit, soit $\Psi(1, 2, ..., N)$ la fonction d'onde d'un tel système où i dénote l'ensemble des degrés de liberté (spin et coordonnées d'espace au sens large) ; notons que i n'est pas le numéro d'une particule, mais le numéro d'un point de l'espace-spin pour une particule. Le postulat ci-dessus affirme que les seuls états possibles, suivant la nature des particules composant le système, sont :

- les états symétriques tels que :

$$\boxed{\Psi(..., i, ..., j, ...) = +\,\Psi(..., j, ..., i, ...)} \tag{22.6}$$

  que l'on notera Ψ_{S} et appartiennent à un sous-espace $\mathcal{E}_{\mathrm{S}}$ de l'espace des états $\mathcal{E}$;

- les états antisymétriques tels que :

$$\boxed{\Psi(..., i, ..., j, ...) = -\,\Psi(..., j, ..., i, ...)} \tag{22.7}$$

  que l'on notera Ψ_{A} et appartiennent à un sous-espace $\mathcal{E}_{\mathrm{A}}$ de l'espace des états $\mathcal{E}$.

On appelle *bosons* les particules pour lesquelles la fonction d'onde doit être symétrique, *fermions* celles pour lesquelles la fonction d'onde doit être antisymétrique. Toutes les particules *élémentaires* connues à ce jour vérifient en outre la règle empirique suivante :

- les particules de spin entier sont des bosons, et sont donc décrits par des états symétriques satisfaisant (22.6) ;
- les particules de spin demi-entier sont des fermions, et sont donc décrits par des états antisymétriques satisfaisant (22.7).

Selon ce que l'on sait actuellement, la matière est faite de fermions, tandis que les bosons se chargent de médier les interactions. Dès que ce postulat est admis pour les particules considérées comme élémentaires, il s'applique également aux particules composites. En effet, échanger deux telles particules complexes revient à échanger simultanément toutes les particules composant la première avec toutes les particules composant la deuxième. L'état est inaltéré si les constituants élémentaires sont des bosons (aucun changement de signe) ou s'il s'agit de fermions en nombre pair, un nombre pair de fermions donnant

un spin entier donc un spin de boson. À l'inverse, des particules complexes formées d'un nombre impair de fermions sont des fermions[9] (spin total demi-entier) et l'échange de deux telles particules par l'échange de leurs constituants élémentaires donnera globalement un signe $-$.

L'association stricte entre la symétrie de permutation et le caractère entier ou demi-entier du spin – énoncée ci-dessus en tant que règle empirique – constitue en fait le Théorème spin-statistique (*Spin-statistics theorem*), énoncé par Pauli en 1940 [65], et reformulé par Schwinger[10] [166] – pour une discussion récente sur cette question, voir [66].

Il est essentiel de remarquer que, dans tous les cas, l'échange de deux particules identiques laisse toujours invariant le *module carré* de la fonction d'onde. De fait, cette opération ne doit pas modifier l'information physique contenue dans celle-ci, laquelle s'exprime comme des probabilités ; de cette invariance résulte l'égalité des modules, ce qui laisse la place à une phase globale constante. *A priori*, il faut donc :

$$\Psi(..., i, ..., j, ...) = \mathrm{e}^{\mathrm{i}\alpha}\, \Psi(..., j, ..., i, ...) \ . \tag{22.8}$$

Clairement, l'affirmation du postulat de symétrisation est plus forte que celle traduisant l'invariance nécessaire des probabilités : ce postulat dit que la phase α ne peut être *que* 0 (bosons) *ou* π (fermions).

Il est facile de voir que la symétrie de permutation de l'état d'un système donné ne peut pas changer au cours du temps. Remarquons d'abord que pour un système de N particules identiques, le Hamiltonien est inaltéré si on échange deux ou plusieurs particules : étant identiques, elles figurent exactement de la même façon dans H. Par ailleurs, supposons-les en interaction deux à deux par un certain opérateur $W(r_{ij})$ où $r_{ij} = \|\vec{r}_i - \vec{r}_j\|$; $W(r_{ij})$ est par exemple l'interaction de Coulomb $\frac{e'^2}{r_{ij}}$ entre deux électrons. Toutes les particules ont notamment la même masse m par hypothèse, et sont en général soumises individuellement à un champ de force donnant à chacune l'énergie potentielle $V(\vec{r}_i)$. Dans ces conditions, le Hamiltonien s'écrit :

$$H = \sum_{i=1}^{N} \left[\frac{\vec{p}_i^{\,2}}{2m} + V(\vec{r}_i)\right] + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1,\, j\neq i}^{N} W(r_{ij}) \ ; \tag{22.9}$$

le facteur 1/2 évite de compter deux fois l'interaction entre deux particules données[11]. Le terme cinétique est invariant quand on change l'ordre dans la sommation ; quant au

[9] C'est pourquoi les deux isotopes de l'hélium ^{3}He et ^{4}He ont des propriétés radicalement différentes. ^{3}He est un fermion, alors que ^{4}He est un boson ; la superfluidité de ce dernier à basse température est la manifestation de la condensation de Bose conventionnelle.

[10] Julian Seymour Schwinger (1918-1994) a partagé le Nobel en 1965 avec Feynman et Tomonaga pour leurs travaux sur l'Électrodynamique quantique.

[11] De ce point de vue, le terme à deux corps dans (22.9) s'écrit tout autant :

$$\sum_{i=1}^{N} \sum_{j=i+1}^{N} W(r_{ij}) \ . \tag{22.10}$$

terme d'énergie potentielle, il est également inaltéré dans toute permutation des indices i puique, précisément, les particules ont toutes les mêmes attributs propres (leur charge, par exemple) : H est bien symétrique dans toute permutation des indices i, j,....

Il en va d'ailleurs de même pour toute autre observable de ce système : s'il n'en était pas ainsi, la mesure de cette observable permettrait justement de distinguer les particules or celles-ci sont précisément réputées identiques[12]. Par exemple, si $\vec{S}_i$ désigne le spin de l'une quelconque des particules d'un système de N particules identiques, le spin total est :

$$\vec{S} = \sum_{i=1}^{N} \vec{S}_i \tag{22.11}$$

et est manifestement invariant dans toute permutation (on peut additionner les spins dans n'importe quel ordre sans changer le résultat).

Montrons maintenant que la propriété de symétrie de permutation (symétrie ou antisymétrie) se conserve au cours du temps, une nécessité pour ne pas obtenir, le temps passant, une contradiction entre les postulats généraux énoncés antérieurement et le postulat de symétrisation – étant entendu que toute observable pouvant faire l'objet d'une mesure est nécessairement totalement symétrique. Chaque espace $\mathcal{E}_\text{A}$ ou $\mathcal{E}_\text{S}$, où se trouve le vecteur d'état d'un système à un instant donné, est ainsi *stable* par l'évolution dans le temps et forme un sous-espace hermétiquement clos dont le système ne peut pas sortir.

Soit un état initial $\Psi(1, 2, ..., N\,;\, t)$ de symétrie donnée (symétrique ou antisymétrique). Par l'équation de Schrödinger, on a :

$$\Psi(1, 2, ..., N\,;\, t + \mathrm{d}t) = \Psi(1, 2, ..., N\,;\, t) + \frac{\mathrm{d}t}{\mathrm{i}\hbar} H(1, ..., N)\Psi(1, 2, ..., N\,;\, t) + \mathcal{O}(\mathrm{d}t^2)\ . \tag{22.12}$$

Effectuons l'échange de deux quelconques des particules dans les deux membres. Il vient :

$$\begin{aligned}\Psi(..., j, ..., i, ...;\, t + \mathrm{d}t) &= \Psi(..., j, ..., i, ...;\, t) + \\ &\frac{\mathrm{d}t}{\mathrm{i}\hbar} H(1, ..., N)\Psi(..., j, ..., i, ...;\, t) + \mathcal{O}(\mathrm{d}t^2)\end{aligned} \tag{22.13}$$

(H ne change pas). Si $\Psi(...\,;\, t)$ est une fonction paire (symétrique) Ψ_S, le second membre est inaltéré, égal à $\Psi_\text{S}(..., i, ..., j, ...\,;\, t + \mathrm{d}t)$, d'où :

$$\Psi_\text{S}(..., j, ..., i, ...;\, t + \mathrm{d}t) = \Psi_\text{S}(...i, ..., j, ...;\, t + \mathrm{d}t) \tag{22.14}$$

et la fonction d'onde à l'instant $t + \mathrm{d}t$ est encore paire. Au contraire, si $\Psi(...\,;\, t)$ est une fonction impaire (antisymétrique) Ψ_A, on a maintenant :

$$\begin{aligned}\Psi_\text{A}(..., j, ..., i, ...\,;\, t + \mathrm{d}t) &= -\Psi_\text{A}(..., i, ..., j, ...;\, t) + \\ &\frac{\mathrm{d}t}{\mathrm{i}\hbar} H(1, ..., N)(-1)\Psi_\text{A}(...i, ..., j, ...\,;\, t) + \mathcal{O}(\mathrm{d}t^2)\ ,\end{aligned} \tag{22.15}$$

[12]Notons que la réciproque est fausse : pour deux particules non-identiques, le Hamiltonien peut néanmoins être symétrique. Ainsi, le Hamiltonien du couple électron-positron en interaction de Coulomb est symétrique, et pourtant l'électron et le positron ne sont pas identiques (charges opposées).

puisque les deux fonctions au second membre changent de signe par hypothèse mais pas H. Il en résulte :

$$\Psi_{\mathrm{A}}(..., j, ..., i, ... ; t + \mathrm{d}t) = -\Psi_{\mathrm{A}}(...i, ..., j, ... ; t + \mathrm{d}t) . \tag{22.16}$$

Ainsi, de proche en proche, le caractère de symétrie de l'état initial se perpétue au cours du temps : c'est ce que l'on appelle parfois une règle de supersélection. Ceci a des conséquences importantes : par exemple, un système de bosons ne peut évoluer en se fragmentant en un nombre impair de fermions. Tout morcellement ne peut donc se faire que par émission de fermions par paires de façon à préserver le caractère bosonique de l'état. De même, un système de fermions ne peut créer un boson *libre* que si celui-ci est formé de deux fermions. D'un autre côté, deux fermions peuvent échanger des bosons virtuels, particules véhiculant l'interaction entre les deux fermions.

♦ *Remarques importantes*

1. La propriété de symétrie dans l'échange se réfère essentiellement à la fonction d'onde *complète*, impliquant des degrés de liberté orbitaux ("coordonnées", c'est-à-dire coordonnées ou impulsions suivant la représentation choisie) et ceux de spin. Dans l'exemple des deux spins 1/2 traité plus haut – et ne considérant toujours que les variables de spin –, on verra que le postulat de symétrisation permet de retenir, pour la partie de la fonction d'onde ne dépendant que du spin, deux combinaisons linéaires et deux seulement, définies comme suit (à une phase près)[13] :

   (a) $a = +1$ et $b = +1$ correspondant à la composante $M_S = 0$ de l'état triplet (spin total $S = 1$) et donnant la combinaison :

   $$|\chi_1\rangle = \frac{1}{\sqrt{2}}\Big(|+\frac{1}{2}, -\frac{1}{2}\rangle + |-\frac{1}{2}, +\frac{1}{2}\rangle\Big) . \tag{22.17}$$

   (b) $a = +1$ et $b = -1$ correspondant à l'état singulet (spin total $S = 0$) et donnant la combinaison :

   $$|\chi_0\rangle = \frac{1}{\sqrt{2}}\Big(|+\frac{1}{2}, -\frac{1}{2}\rangle - |-\frac{1}{2}, +\frac{1}{2}\rangle\Big) . \tag{22.18}$$

   La fonction de spin de l'état singulet est antisymétrique dans l'échange, celle de l'état triplet est au contraire symétrique – et pourtant il s'agit dans tous les cas d'un système de deux fermions dont l'état global doit être antisymétrique. Il n'y a pas de difficulté : les vecteurs $|\chi_0\rangle$ et $|\chi_1\rangle$ ne prennent pas en compte les degrés de liberté orbitaux, il s'agit seulement de ce que l'on appelle parfois *l'état de spin*, qui n'est que l'un des *morceaux* de l'état total (pour N quelconque, un tel état de spin peut en fait avoir une symétrie[14] très différente : elle n'est pas quelconque mais entièrement déterminée par la valeur de M_S, et peut s'obtenir à l'aide de diagrammes appelés

[13] Ces deux états coïncident avec les états intriqués de Bell introduits au chapitre 21, éq.(21.198).

[14] En outre, dans le cas où le Hamiltonien ne dépend pas du spin, on pourra toujours construire la fonction à symétriser conformément au postulat à partir du *produit* d'une fonction d'espace par une

tableaux d'Young [15]). La probabilité $P(a, b)$ (22.4) est maintenant déterminée : elle vaut 1 pour l'état triplet, 0 pour l'état singulet, le postulat de symétrisation ayant bien supprimé la difficulté ; la description doit d'ailleurs être complétée, justement pour tenir compte des degrés orbitaux.

2. Pour des fermions, la fonction d'onde change de signe à chaque échange de deux particules – ce que l'on appelle une transposition dans la théorie du groupe des permutations (voir section suivante). Si l'on fait deux échanges, la fonction d'onde au total ne change pas de signe. Par exemple :

$$\Psi_{\mathrm{A}}(1,\,2,\,3,\,4) = (-1)^2\,\Psi_{\mathrm{A}}(2,\,1,\,4,\,3)\ . \tag{22.19}$$

Plus généralement, pour une permutation quelconque, Ψ acquiert un signe qui est, par définition, la *signature* de la permutation (voir ci-dessous, section 22.3).

3. On rappellera ci-dessous qu'une permutation quelconque est factorisable en un produit de transpositions – ce qui permet de dire que les transpositions sont les *atomes* des permutations. Ceci étant réalisé, on voit qu'il faut, et il suffit, de dire ce que fait une fonction d'onde sous l'action d'une transposition, quelle que soit cette dernière. Pour des fermions, le postulat d'antisymétrisation affirme que toute transposition change le signe de la fonction d'onde.

4. On rencontre parfois l'argument suivant. Considérons le cas $N = 2$; soit $f(1, 2)$ une fonction quelconque, et T la transposition qui échange 1 et 2 :

$$Tf(1,\,2) \stackrel{\text{déf}}{=} f(2,\,1)\ ; \tag{22.20}$$

si on effectue à nouveau la transposition sur la fonction de droite, il vient :

$$Tf(2,\,1) = f(1,\,2)\ , \tag{22.21}$$

d'où l'on déduit $T^2 = \mathbf{1}$ – ce qui n'est pas un *scoop.* (22.21) montre aussi que les valeurs propres de T sont ± 1, ce qui est indéniable, et que les fonctions propres de T, $\phi(1, 2)$, sont soit symétriques, soit antisymétriques puisqu'elles satisfont :

$$T\phi(1,\,2) = (\pm 1)\phi(1,\,2)\ . \tag{22.22}$$

Effectuons maintenant ces opérations avec une fonction d'onde $\Psi(1, 2)$; on est alors tenté d'écrire :

$$T\Psi(1,\,2) = (\pm 1)\Psi(1,\,2)\ , \tag{22.23}$$

---

fonction de spin. La symétrisation détruit en général la séparation des variables espace-spin : après cette opération, la fonction d'onde n'est plus *un* simple produit de deux facteurs, l'un dépendant des coordonnées seules, l'autre des variables de spin, mais est une combinaison linéaire de tels produits ; ceux-ci s'échangent mutuellement avec les bons signes lors d'une permutation, en sorte que la fonction d'onde totale a la symétrie requise. La séparation en un seul produit ne subsiste que dans deux cas : pour $N = 2$ d'une part, et pour les composantes $M_S = \pm S$ de l'état de spin maximum $S = N/2$ pour N fermions de spin 1/2. Dans tous les autres cas, la fonction d'onde est une combinaison linéaire de produits : c'est un exemple d'*intrication*, ici entre degrés de liberté d'espace et degrés de liberté de spin.

ce qui tendrait à montrer que l'on peut se passer du postulat d'antisymétrisation – cette conclusion imposant l'idée que le raisonnement a une faille, ou n'est pas pertinent pour les états physiquement acceptables.

En fait, le postulat est bien nécessaire : imposer (22.23) à $\Psi(1,\,2)$, c'est contraindre les états physiques à être propres de T, or rien ne l'exige : la seule nécessité est que Ψ et $T\Psi$ représentent le *même état physique*. Or si effectivement le carré de T est l'identité (l'opération qui ne fait rien), ceci n'exclut pas pour autant que l'on ait :

$$T^2\Psi(1,\,2) = \mathrm{e}^{\mathrm{i}\alpha}\Psi(1,\,2)\ , \tag{22.24}$$

α étant une phase quelconque. En effet, T^2 ne doit changer aucune prévision physique avec Ψ, mais comme cette fonction et $\mathrm{e}^{\mathrm{i}\alpha}\Psi$ représentent un seul et même état[15], il n'y aurait *a priori* aucune contre-indication à prendre α quelconque : en toute rigueur, un état quantique n'est pas représenté par un vecteur mais par un *rayon*, c'est-à-dire par un vecteur *à une phase près*, l'ensemble des rayons formant un *espace de Hilbert projectif*. L'argument tendant à prouver que les deux valeurs $\alpha = 0$ *ou* π sont une simple *conséquence* de $T^2 = \mathbf{1}$ est spécieux : on a bel et bien besoin du postulat d'antisymétrisation.

5. L'affirmation du 7ᵉ postulat concerne exclusivement les particules *élémentaires*, et permet de construire les états de bonne symétrie décrivant les particules *composites*. Avec ce jeu de construction, on n'échappe cependant pas à la règle : il n'existe que des bosons *ou* des fermions puisque, par addition des moments cinétiques entiers ou demi-entiers des particules élémentaires formant un objet complexe dûment identifié, on ne peut obtenir qu'un entier ou un demi-entier. C'est ainsi que deux électrons formant une paire liée du genre paire de Cooper constituent un boson. Quand on tient compte du spin nucléaire $\vec{I}$, l'atome d'hydrogène (normal) est un boson puisque le proton (tout comme le neutron) a un spin $1/2$; avec le spin électronique $\vec{S}$, la somme $\vec{F} = \vec{I}+\vec{S}$ peut prendre les valeurs $f = 0,\,1$ (l'introduction du moment orbital $\vec{L}$ de l'électron ne change rien, puisque l est forcément entier).

   Assez récemment ([167], [168]), il a été suggéré que, pour des systèmes de dimension $D = 2$, pourraient exister des objets n'appartenant à aucune des deux classes bosons ou fermions. Dans la formulation générale de l'invariance par permutation du contenu physique de la fonction d'onde, il n'a été nulle part question de la dimension D de l'espace $\mathbb{R}^D$ où "vit" le système ; afin de sentir le rôle que celle-ci peut jouer, essayons de visualiser géométriquement la notion de permutation, plus précisément (et c'est suffisant), celle de transposition.

   Soit donc à permuter deux objets notés 1 et 2, une opération que l'on peut représenter par un arc reliant 1 à 2 et 2 à 1 (voir fig. 22.2) : 1 prend la place de 2, 2 celle de 1. Si l'on recommence la même opération, 1 et 2 se retrouvent à leurs places initiales. Cette double opération est ainsi représentée par la courbe fermée Γ de la figure. Il est naturel de penser que ces trajets induisent un facteur de phase

[15] La phase α est ici bien sûr supposée constante. Une phase variable (fonction de la coordonnée) se rencontre par exemple lorsque l'on effectue une transformation de jauge, voir chapitre 17, section 17.4, notamment éq. (17.150).

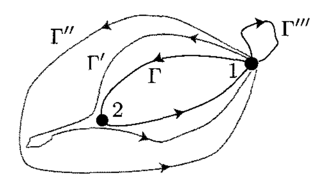


Figure 22.2: Illustration géométrique d'une permutation élémentaire (transposition) effectuée deux fois.

pour la fonction d'onde de 1 et 2, la phase étant donnée par une certaine intégrale de ligne calculée sur le chemin effectivement suivi. Ceci admis, cette phase ne saurait dépendre en fait du chemin parcouru, que l'on peut finalement tracer un peu comme on veut : la *même* phase apparaît donc pour les boucles Γ, Γ' et Γ''.

Quand les deux objets 1 et 2 sont élémentaires, et situés dans l'espace $\mathbb{R}^3$, la boucle peut toujours être déformée en un chemin du genre Γ''', lequelle peut visiblement être réduite à un point, auquel cas la phase est nulle, en conformité avec ce qu'affirme le postulat de symétrisation : quand on a fait deux fois la même transposition, le facteur de phase est $(\pm 1)^2$ selon qu'il s'agit de bosons ou de fermions (la phase vaut 0 ou π respectivement), facteur qui vaut toujours $+1$. À trois dimensions, on peut toujours éviter un point en passant par au-dessus ou par au-dessous.

Au contraire, quand la liberté de mouvement est strictement confinée au plan $\mathbb{R}^2$, les courbes Γ, Γ' et Γ'' appartiennent à une certaine classe d'équivalence géométrique (elles *entourent* toutes le point 2), et la courbe Γ''' à une autre classe, celle des boucles qui laissent 2 à l'extérieur[16]. Dans $\mathbb{R}^2$, ces deux classes ne sont pas réductibles l'une à l'autre : là réside la particularité de la dimension $D = 2$.

Ainsi, quand on admet la représentation géométrique ci-dessus, il apparaît clairement que la dimension $D = 2$ possède une spécificité qui la rend singulière. De fait, il a été suggéré ([167], [168]) que, pour des objets *composites* définis dans un espace à deux dimensions, la phase associée à une transposition pourrait prendre d'autres valeurs que 0 et π ; ces objets ont été appelés *anyons*, contraction de *any* et du *on* générique favori des physiciens pour désigner une "particule", qu'elle soit élémentaire ou non. Ces *anyons* ne sont pas des excitations (particules) élémentaires, ce qui infirmerait le postulat de symétrisation, mais sont des *molécules* d'objets élémentaires, formé au moins d'une certaine paire. La particularité est que la fonction d'onde de cette paire se transforme comme suit :

$$T_{12}\Psi(1,\,2) \stackrel{\text{déf}}{=} \Psi(2,\,1) = \mathrm{e}^{\mathrm{i}\theta}\Psi(1,\,2) \ , \qquad (22.25)$$

où la phase θ peut être différente de 0 et π. Ceci admis, un *anyon* n'est ni un boson, ni un fermion, mais quelque chose entre les deux, dont l'existence n'est possible que dans un espace bidimensionnel.

[16]L'analogie qui vient immédiatement à l'esprit est la situation conduisant au théorème des résidus pour une fonction holomorphe dans un domaine non simplement connexe : l'analogue du point 2 est une singularité de la fonction (un pôle par exemple). L'intégrale sur un cycle de la fonction est nulle tant que le cycle, quel qu'il soit, peut être réduit à un point (c'est Γ''') ; en revanche, toute boucle entourant la singularité donne une intégrale non-nulle. La richesse du *plan* complexe repose sur sa dimension spatiale.

Cette particularité dimensionnelle pourrait faire croire que ces objets exotiques n'ont qu'un intérêt académique, la dimension "naturelle" des systèmes physiques étant l'espace $\mathbb{R}^3$. En fait, si effectivement tous les systèmes sont plongés dans $\mathbb{R}^3$, certains, lamellaires, possèdent des propriétés d'anisotropie si marquées qu'ils sont modélisables par des systèmes de dimensionnalité inférieure – le graphite, du point de vue de la conductivité, en fournit un exemple banal, où les électrons vivent pratiquement à deux dimensions, précisément (dans les plans de graphite).

Dans les composés bidimensionnels, à très basse température (quelques K), l'effet Hall présente des particularités remarquables, puisque la résistance Hall, au lieu de croître *linéairement* en fonction du champ magnétique appliqué, a la forme d'un escalier dont les marches surviennent à des valeurs *entières* d'une certaine quantité, appelée[17] *facteur de remplissage.* Par la suite, on a découvert qu'il existe aussi des marches pour les valeurs non-entières, toutes caractérisables par des fractions simples (1/3, 1/5, 2/5,...), un effet *anormal* que l'on a appelé effet Hall *fractionnaire.* Laughlin[18] [169] a proposé un cadre théorique pour expliquer cet effet, qui introduit des excitations de type anyonique, ayant par ailleurs une charge non-entière. Sur le plan expérimental, des charges fractionnaires $e/3$ ont été observées [170] dans un système bidimensionnel soumis à un fort champ magnétique perpendiculaire ; en ce qui concerne plus particulièrement les anyons, une expérience [171] prétend à ce jour apporter une preuve de leur existence et de la statistique *fractionnaire* qui les caractérise, expérience dont l'interprétation est d'ailleurs controversée [172]. ♦

22.3 Permutations. Opérateurs de symétrisation et d'antisymétrisation

Cette section présente les aspects utiles du groupe des permutations, un groupe aux propriétés nombreuses et parfois complexes (pour en savoir plus, voir l'ouvrage classique de Rutherford [173]) ; soit $\mathcal{S}_N \equiv \{P_\lambda\}_\lambda$ l'ensemble des permutations de N objets. Une permutation peut se noter :

$$\begin{pmatrix} 1 & 2 & \ldots & N \\ i_1 & i_2 & \ldots & i_n \end{pmatrix} , \tag{22.26}$$

la deuxième ligne donnant le résultat de la permutation des N premiers entiers écrits dans l'ordre naturel dans la première ligne. Cet ensemble est visiblement un groupe ; le produit de deux permutations est une permutation, que l'on détermine comme suit :

$$\begin{pmatrix} 1 & 2 & \ldots & N \\ i_1 & i_2 & \ldots & i_N \end{pmatrix}\begin{pmatrix} 1 & 2 & \ldots & N \\ j_1 & j_2 & \ldots & j_N \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & \ldots & j_N \\ k_1 & k_2 & \ldots & k_N \end{pmatrix}\begin{pmatrix} 1 & 2 & \ldots & N \\ j_1 & j_2 & \ldots & j_N \end{pmatrix} . \tag{22.27}$$

[17]Comme on le verra dans le chapitre 30, l'existence de la symétrie discrète de translation d'un cristal (parfait) conduit, dans une théorie à une particule, à l'existence de *bandes d'énergie* pour les électrons. À température nulle, l'état fondamental s'obtient en remplissant ces bandes conformément au Principe de Pauli. Une fois épuisé le stock d'électrons à caser, une bande électronique peut se trouver plus ou moins pleine ; par définition, le facteur de remplissage d'une bande est la fraction d'états occupés dans cette bande.

[18]Robert Betts Laughlin a partagé le Nobel 1998 avec Horst L. Störmer et Daniel C. Tsui *"for their discovery of a new form of quantum fluid with fractionally charged excitations"*.

Le produit au second membre n'est autre que :

$$\begin{pmatrix} 1 & 2 & \dots & N \\ k_1 & k_2 & \dots & k_N \end{pmatrix} . \tag{22.28}$$

L'élément neutre est la permutation identité $i_k = k$ quelque soit k ; manifestement, tout élément a un inverse. Il y a $N!$ éléments dans $\mathcal{S}_N \equiv \{P_\lambda\}_{\lambda=1,2,\dots,N!}$, qui est appelé *groupe symétrique*. Les permutations P_λ agissent sur les fonctions d'onde comme suit ; si :

$$P_\lambda = \begin{pmatrix} 1 & 2 & \dots & N \\ i_1 & i_2 & \dots & i_N \end{pmatrix} , \tag{22.29}$$

alors :

$$\boxed{P_\lambda \Psi(1, 2, ..., N) \stackrel{\text{déf}}{=} \Psi(i_1, i_2, ..., i_N)} \tag{22.30}$$

Les P_λ sont visiblement des opérateurs unitaires ; en effet, un produit scalaire $(P_\lambda\Psi,\, P_\lambda\Phi)$ est une intégration sur les variables d'espace continues et une sommation sur les variables de spin discrètes ; toutes ces sommations portent sur des variables muettes et le résultat ne dépend en aucune façon de toute permutation effectuée sur *l'ensemble* de ces variables. Il vient donc :

$$(P_\lambda\Psi,\, P_\lambda\Phi) = (\Psi,\, \Phi) \quad \forall\, \lambda , \tag{22.31}$$

relation caractéristique d'un opérateur unitaire. Toute permutation étant unitaire, son adjoint est égal à son inverse :

$$\boxed{P_\lambda^\dagger = P_\lambda^{-1} \quad \forall\, \lambda} \tag{22.32}$$

Une permutation n'étant pas égale à son inverse en général, P_λ n'est pas hermitique – sauf dans le cas des permutations particulières appelées transpositions, voir ci-dessous.

Un lemme sera utile, dit *lemme de réarrangement*. Une permutation P_μ étant donnée, tous les produits $P_\mu P_\lambda$ où λ décrit le groupe *entier*, donnent une fois et une seule les $N!$ permutations. En effet, par définition de la structure de groupe, le produit $P_\mu P_\lambda$ est une certaine permutation P_ν. Quand on effectue ces $N!$ produits, si on trouvait deux fois la même permutation P_ν, cela signifierait qu'il existe λ_1 et λ_2 tels que :

$$P_\mu P_{\lambda_1} = P_\nu \ , \quad P_\mu P_{\lambda_2} = P_\nu \quad \Longleftrightarrow \quad P_{\lambda_1} = P_\mu^{-1} P_\nu \ , \quad P_{\lambda_2} = P_\mu^{-1} P_\nu \tag{22.33}$$

d'où $P_{\lambda_1} = P_{\lambda_2}$. Le lemme de réarrangement peut s'écrire formellement :

$$\boxed{P_\lambda \mathcal{S}_N = \mathcal{S}_N P_\lambda = \mathcal{S}_N \quad \forall\, P_\lambda} \tag{22.34}$$

Certaines permutations jouent un rôle particulièrement important, ce sont les transpositions. Une transposition échange deux objets seulement ; notant T_{ij} la permutation qui se borne à permuter i et j, on a donc par définition :

$$T_{ij} \stackrel{\text{déf}}{=} \begin{pmatrix} 1 & 2 & 3 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & N \\ 1 & 2 & 3 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & N \end{pmatrix} , \tag{22.35}$$

que l'on note plus simplement :

$$T_{ij} = \begin{pmatrix} i & j \\ j & i \end{pmatrix} . \tag{22.36}$$

Comme toute permutation, les transpositions sont unitaires, mais comme leur carré est l'identité, elles sont également hermitiques :

$$\boxed{T_{ij}^2 = \mathbf{1} \quad \Longleftrightarrow \quad T_{ij} = T_{ij}^{-1} = T_{ij}^\dagger} \tag{22.37}$$

Enfin, évidemment, $T_{ij} = T_{ji}$.

Un résultat important est le suivant : toute permutation peut se décomposer en produit de transpositions, la décomposition n'étant pas unique (on peut en écrire autant qu'on veut[19]) mais, pour une permutation donnée P_λ le nombre σ_λ de transpositions a une *parité* parfaitement définie, que l'on appelle précisément la *parité* de la permutation ; le nombre $(-1)^{\sigma_\lambda} = \pm 1$ est appelé *signature de la permutation* – la signature d'une transposition est égale à -1. Par exemple, la permutation circulaire :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \tag{22.38}$$

peut se décomposer comme suit :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = T_{12}T_{23} . \tag{22.39}$$

Pour cette permutation, σ vaut 2 (ou n'importe quel nombre pair), il s'agit donc d'une permutation paire. Les deux permutations circulaires, pour $N = 3$, ont pour signature $(-1)^2 = +1$.

La signature de la permutation P_λ, $(-1)^{\sigma_\lambda}$, permet d'exprimer simplement le résultat d'une permutation quelconque agissant sur une fonction d'onde de fermions, puisque chaque transposition introduit un facteur (-1) :

$$\boxed{P_\lambda \Psi_\mathrm{A}(1, 2, ..., N) = (-1)^{\sigma_\lambda}\Psi_\mathrm{A}(1, 2, ..., N) \qquad \text{(fermions)}} \tag{22.40}$$

Evidemment, pour des bosons, on a :

$$\boxed{P_\lambda \Psi_\mathrm{S}(1, 2, ..., N) = \Psi_\mathrm{S}(1, 2, ..., N) \qquad \text{(bosons)}} \tag{22.41}$$

Ces définitions étant données, on peut maintenant montrer comment construire les fonctions physiquement acceptables, c'est-à-dire celles ayant la symétrie requise vis-à-vis des permutations. L'idée est d'engendrer, partant d'une fonction Φ donnée quelconque, la bonne fonction Ψ_A ou Ψ_S, suivant la nature du système considéré.

[19] Une certaine factorisation étant trouvée, on peut insérer n'importe où, et autant de fois que l'on veut, le produit $T_{ij}T_{ji}$, ce qui augmente d'un nombre pair le nombre des facteurs.

Avant de donner la méthode générale, l'idée centrale peut être saisie en considérant le cas de $N = 2$ particules. Une fonction $\Phi(1,2)$ étant choisie, il est aisé d'en déduire les fonctions ayant la bonne symétrie ; ainsi, la combinaison linéaire $\Phi(1,2) + \Phi(2,1)$ est symétrique dans l'échange, en revanche $\Phi(1,2) - \Phi(2,1)$ est antisymétrique ; ces deux fonctions sont donc respectivement du type Ψ_{S} et Ψ_{A}, et peuvent s'écrire[20] :

$$\Psi_{\text{S}} = (\mathbf{1} + T_{12})\,\Phi(1,2) \ , \qquad \Psi_{\text{A}} = (\mathbf{1} - T_{12})\,\Phi(1,2) \tag{22.42}$$

On voit ainsi qu'en faisant agir sur une fonction quelconque la bonne combinaison linéaire d'opérateurs de permutation, on obtient la fonction ayant la symétrie voulue. C'est cette idée que l'on généralise maintenant, en définissant les deux opérateurs[21] Y_{S} et Y_{A} :

$$\boxed{Y_{\text{S}} \stackrel{\text{déf}}{=} \frac{1}{N!}\sum_{\lambda=1}^{N!} P_\lambda \qquad Y_{\text{A}} \stackrel{\text{déf}}{=} \frac{1}{N!}\sum_{\lambda=1}^{N!} (-1)^{\sigma_\lambda}\, P_\lambda} \tag{22.43}$$

que l'on peut collectivement dénoter Y_{I} :

$$Y_{\text{I}} = \frac{1}{N!}\sum_{\lambda=1}^{N!} \varepsilon_\lambda\, P_\lambda \ , \qquad \varepsilon_\lambda = \begin{cases} 1 & (\text{I} = \text{S, bosons}) \\ (-1)^{\sigma_\lambda} & (\text{I} = \text{A, fermions}) \end{cases} \ . \tag{22.44}$$

Les Y_{I} sont visiblement des opérateurs hermitiques puisque chaque P_λ est unitaire : dans la sommation dont on a pris l'adjoint, $P_\lambda^\dagger$ peut être remplacé par P_λ^{-1}, qui est un autre terme de la même sommation, avec évidemment[22] le même ε_λ.

On va maintenant montrer que $[Y_{\text{S}},\, Y_{\text{A}}] = 0$, afin d'en déduire commodément :

$$\boxed{Y_{\text{S}} Y_{\text{A}} = Y_{\text{A}} Y_{\text{S}} = 0} \tag{22.45}$$

Notons d'abord que :

$$P_\nu = P_\lambda P_\mu \implies \sigma_\nu = \sigma_\lambda + \sigma_\mu \iff \varepsilon_\nu = \varepsilon_\lambda \varepsilon_\mu \iff \varepsilon_\lambda = \varepsilon_\mu \varepsilon_\nu \ , \tag{22.46}$$

par définition des σ_λ : le nombre de transpositions en lesquelles se décompose P_ν est la somme des deux nombres correspondants pour P_λ et P_μ ; la dernière égalité dans (22.46) s'obtient en multipliant la précédente membre à membre par ε_μ et sachant que $\varepsilon_\mu^2 = 1$ dans tous les cas. Maintenant :

$$Y_{\text{A}} Y_{\text{S}} = \frac{1}{N!^2}\sum_{\lambda,\,\mu=1}^{N!} \varepsilon_\lambda\, P_\lambda P_\mu = \frac{1}{N!^2}\sum_{\nu,\,\mu=1}^{N!} \varepsilon_\mu \varepsilon_\nu\, P_\nu = \frac{1}{N!^2}\Big(\sum_{\mu=1}^{N!} \varepsilon_\mu\Big)\sum_{\nu=1}^{N!} \varepsilon_\nu\, P_\nu \ . \tag{22.47}$$

[20] Ces combinaisons linéaires ne sont pas normalisées.

[21] La notation Y est un clin d'œil à Alfred Young (1873-1940), mathématicien anglais surtout connu pour ses travaux sur le groupe des permutations. Il a notamment introduit les *tableaux d'Young*, qui constituent une visualisation très commode des représentations irréductibles de $\mathcal{S}_N$. La plupart de ses articles furent rédigés après qu'il fut devenu pasteur près de Cambridge, où il recommença à enseigner après 1926. Concernant les tableaux (diagrammes) d'Young, on peut en trouver une introduction très pédagogique dans l'ouvrage de Messiah [15], et plus approfondie dans celui de Rutherford [173].
Le facteur $\frac{1}{N!}$ est introduit pour l'idempotence – voir (22.53).

[22] La factorisation de P_λ^{-1} s'obtient en renversant l'ordre des facteurs dans celle de P_λ : ceci ne change pas le nombre de facteurs !

On trouve visiblement la même expression en effectuant le produit dans l'autre sens – quoique $\mathcal{S}_N$ n'est pas abélien –, d'où $Y_\mathrm{S} Y_\mathrm{A} = Y_\mathrm{A} Y_\mathrm{S}$, soit $[Y_\mathrm{S},\, Y_\mathrm{A}] = 0$.

Maintenant, soit T une transposition quelconque ; tout produit TP_λ donne une autre permutation du groupe (une fois et une seule quand λ prend toutes ses valeurs), avec une signature inversée, d'où $TY_\mathrm{A} = -Y_\mathrm{A}$; cela étant, par l'associativité du produit :

$$TY_\mathrm{A} Y_\mathrm{S} = -Y_\mathrm{A} Y_\mathrm{S} = -Y_\mathrm{S} Y_\mathrm{A} \ , \tag{22.48}$$

où la dernière égalité vient de la commutation des Y_I entre eux. En multipliant à nouveau à gauche par T :

$$T^2 Y_\mathrm{A} Y_\mathrm{S} = -TY_\mathrm{S} Y_\mathrm{A} \quad \Longleftrightarrow \quad Y_\mathrm{A} Y_\mathrm{S} = -TY_\mathrm{S} Y_\mathrm{A} \ , \tag{22.49}$$

la dernière égalité venant de $T^2 = \mathbf{1}$; comme $TY_\mathrm{S} = Y_\mathrm{S}$, il vient finalement :

$$Y_\mathrm{A} Y_\mathrm{S} = -Y_\mathrm{S} Y_\mathrm{A} \ , \tag{22.50}$$

ce qui, allié au fait que les Y_I en fait commutent, établit (22.45). Par comparaison avec (22.47), on en déduit aussi que :

$$\boxed{\sum_{\mu=1}^{N!} \varepsilon_\mu = 0} \tag{22.51}$$

De plus, Y_I est un projecteur :

$$Y_\mathrm{I}^2 = \frac{1}{N!^2} \sum_{\lambda=1}^{N!} \sum_{\mu=1}^{N!} \varepsilon_\lambda \varepsilon_\mu P_\lambda P_\mu = \frac{1}{N!^2} \sum_{\nu=1}^{N!} \varepsilon_\nu P_\nu \sum_{\lambda=1}^{N!} 1 \ . \tag{22.52}$$

La dernière forme est obtenue à partir du lemme de réarrangement ; la deuxième sommation donne le simple facteur $N!$, de sorte que :

$$\boxed{Y_\mathrm{I}^2 = Y_\mathrm{I}} \tag{22.53}$$

Y_I est donc idempotent ; comme Y_I est idempotent et hermitique, c'est bien un projecteur.

Calculons maintenant le produit $P_\mu Y_\mathrm{I}$, où P_μ est une permutation quelconque mais donnée. Il vient :

$$P_\mu Y_\mathrm{I} = \frac{1}{N!} \sum_{\lambda=1}^{N!} \varepsilon_\lambda P_\mu P_\lambda = \frac{1}{N!} \sum_{\nu=1}^{N!} \varepsilon_\mu \varepsilon_\nu P_\nu \equiv \varepsilon_\mu Y_\mathrm{I} \ . \tag{22.54}$$

Soit maintenant une fonction quelconque Φ ; le résultat de l'action de Y_I sur Φ est soit une fonction de symétrie I, soit une fonction identiquement nulle. En effet, par exemple avec Y_A, on trouve, compte tenu de (22.54) :

$$\boxed{P_\mu Y_\mathrm{A} \Phi = (-1)^{\sigma_\mu}\, Y_\mathrm{A} \Phi} \tag{22.55}$$

La fonction $Y_\text{A}\Phi$ est donc une fonction antisymétrique dans toute transposition ; de la même façon, on a :

$$\boxed{P_\mu Y_\text{S}\Phi = Y_\text{S}\Phi} \tag{22.56}$$

Ainsi, Y_I projette toute fonction donnée sur le sous-espace $\mathcal{E}_\text{I}$ ayant la symétrie voulue. Cette opération engendre donc la bonne fonction, sauf toutefois dans un cas : celui où la fonction Φ de départ a déjà la symétrie *orthogonale* à celle voulue – ce qu'un peu d'attention doit d'ailleurs permettre de réaliser avant de faire quoi que ce soit. Alors, l'action de Y_I projette un vecteur orthogonal au sous-espace de projection et on trouve zéro[23] (en pratique on voit d'emblée si l'on est dans un tel cas particulier). Notons enfin que deux fonctions de symétrie différente A et S sont orthogonales d'après (22.45).

On remarque que chaque Y_I est finalement une certaine combinaison linéaire des opérateurs du groupe, avec des coefficients dûment choisis suivant la symétrie que l'on veut atteindre. La structure de ces opérateurs est tout à fait générale, et réapparaît à chaque fois que l'on souhaite définir des objets se transformant de façon déterminée par les opérateurs d'un groupe donné (c'est aussi ce qui a été noté à propos des rotations, voir le commentaire p. 702 après l'équation (17.205)). On en retrouvera un autre exemple à propos du groupe des translations d'un réseau, dont la conséquence est le théorème de Bloch (voir chapitre 30, notamment la note 43, p. 1398).

La fonction $Y_\text{I}\Phi$ n'est pas en général normalisée à l'unité, même si Φ l'est ; la condition de normalisation prendra une forme ou une autre selon les propriétés de la foncton Φ elle-même ; les fonctions normalisées de symétrie requise seront notées :

$$\Psi_\text{A} = C_\text{A}\, Y_\text{A}\Phi \ , \qquad \Psi_\text{S} = C_\text{S}\, Y_\text{S}\Phi \ , \tag{22.57}$$

où les C_I sont des constantes de normalisation trouvées en écrivant $\langle\Psi_\text{I}|\Psi_\text{I}\rangle = 1$, soit :

$$1 = \langle\Psi_\text{I}|\Psi_\text{I}\rangle = |C_\text{I}|^2\,\langle Y_\text{I}\,\Phi\,|Y_\text{I}\Phi\rangle = |C_\text{I}|^2\,\langle\,\Phi\,|Y_\text{I}^\dagger Y_\text{I}|\Phi\,\rangle \ ; \tag{22.58}$$

comme Y_I est hermitique et idempotent, la condition de normalisation prend la forme :

$$\boxed{|C_\text{I}|^2\,\langle\,\Phi\,|Y_\text{I}|\Phi\,\rangle = 1} \tag{22.59}$$

Les opérateurs Y_I permettent ainsi de former automatiquement les fonctions ayant la symétrie voulue, à partir d'une fonction Φ choisie sur d'autres critères. Par exemple, Φ peut être sélectionnée en tant que fonction propre d'un ECOC ; dans le cas le plus simple, Φ est seulement propre du Hamiltonien, $H\Phi = E\Phi$. Il est bien clair que l'action de Y_I sur Φ ne modifie pas son statut d'état propre ; en effet, on a d'une part :

$$Y_\text{I}H\Phi = Y_\text{I}E\Phi = EY_\text{I}\Phi \ ; \tag{22.60}$$

d'autre part, H étant invariant dans toute permutation, $[P_\lambda,\, H] = 0\ \forall\,\lambda$, d'où également $[Y_\text{I},\, H] = 0$. Le premier membre de (22.60) est ainsi égal à $HY_\text{I}\Phi$, d'où :

$$H(Y_\text{I}\Phi) = E(Y_\text{I}\Phi) \ , \tag{22.61}$$

[23] Autre façon de voir : si Φ a déjà une symétrie I ou S, alors $Y_\text{I}\Phi = \Phi$. Dès lors $Y_{\text{I}'}\Phi = Y_{\text{I}'}Y_\text{I}\Phi = \delta_{\text{II}'}\Phi$ en vertu de (22.45).

où les parenthèses ne sont là que pour la clarté. (22.61) montre que si Φ est propre de H, alors $Y_{\rm I}\Phi$ est encore propre[24] de H. Le même argument vaut pour toute observable d'un système de N particules identiques, qu'elle appartienne ou non à l'ECOC considéré. En définitive, les projecteurs fabriquent les fonctions de bonne symétrie sans altérer leur éventuel caractère d'état propre requis indépendamment de la symétrie de permutation.

Explicitons le cas $N = 3$. Soit $\Phi(1,2,3)$ une fonction quelconque ; la fonction symétrique engendrée à partir de Φ est :

$$\begin{aligned}\Psi_{\rm S} = C_{\rm S}\frac{1}{3!}\Big[&\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} + \\ &\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\Big]\Phi(1,2,3) \qquad (22.62) \\ = C_{\rm S}\frac{1}{3!}[&\Phi(1,2,3) + \Phi(2,1,3) + \Phi(3,2,1) + \Phi(1,3,2) + \Phi(2,3,1) + \Phi(3,1,2)] \ .\end{aligned}$$

Il est évident que cette dernière fonction est complètement symétrique dans toute permutation. De même, pour le cas antisymétrique :

$$\begin{aligned}\Psi_{\rm A} = C_{\rm A}\frac{1}{3!}\Big[&\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} + \\ &\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\Big]\Phi(1,2,3) \qquad (22.63) \\ = C_{\rm A}\frac{1}{3!}[&\Phi(1,2,3) - \Phi(2,1,3) - \Phi(3,2,1) - \Phi(1,3,2) + \Phi(2,3,1) + \Phi(3,1,2)] \ .\end{aligned}$$

Il est non moins évident que $\Psi_{\rm A}$ change de signe dans toute transposition mais ne change pas si l'on applique l'une des deux permutations circulaires de $\mathcal{S}_3$. D'un autre côté, on voit bien sur cet exemple que si Φ est déjà complètement symétrique, la fonction $\Psi_{\rm A}$ ainsi fabriquée est identiquement nulle : il y a autant de signes $+$ que de signes $-$, conformément à (22.51). Enfin, il est visible que $\Psi_{\rm A}$ et $\Psi_{\rm S}$ sont orthogonales[25].

Afin d'illustrer la différence fondamentale entre bosons et fermions, soit un système formé de deux particules identiques dont l'état est construit à partir d'un simple produit de deux fonctions distinctes ψ et ϕ supposées orthonormalisées. Compte tenu du postulat de symétrisation, les seuls états physiquement acceptables sont de la forme :

$$\Psi(1,2) = C[\mathbf{1} + \varepsilon\, T_{12}]\psi(1)\phi(2) \qquad (22.64)$$

où $\varepsilon = +1$ s'il s'agit de bosons, $\varepsilon = -1$ s'il s'agit de fermions ; comme ψ et ϕ sont orthonormalisées, on a simplement (à une phase près) $C = \frac{1}{\sqrt{2}}$ dans les deux cas. Ceci

[24] D'ailleurs, comme toutes les fonctions $P_\lambda\Phi$ sont propres de H avec la même valeur propre E, *n'importe* quelle combinaison $[\sum_\lambda c_\lambda P_\lambda]\Phi$ reste propre. Les combinaisons résultant des deux projecteurs $Y_{\rm I}$ ne sont que les deux combinaisons particulières satisfaisant à l'exigence du postulat de symétrisation.

[25] Dans le langage de la théorie de la représentation (linéaire) des groupes, on dit que $\Psi_{\rm S}$ et $\Psi_{\rm A}$ se transforment respectivement suivant les représentations irréductibles symétrique et antisymétrique. Deux fonctions se transformant suivant deux représentations irréductibles inéquivalentes sont forcément orthogonales.

peut s'écrire en notation de Dirac :

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left[\mathbf{1} + \varepsilon\, T_{12}\right]|\psi\phi\rangle \tag{22.65}$$

à condition de convenir implicitement que dans le ket (et plus tard dans les bras) les variables sont écrites dans l'ordre naturel de gauche à droite. Soit A une observable à une particule, $A(1,2) = a(1) + a(2)$ à spectre discret :

$$a\,|a_n\rangle = a_n|a_n\rangle \ , \qquad A\,|a_n a_m\rangle = (a_n + a_m)\,|a_n a_m\rangle \ . \tag{22.66}$$

Cherchons maintenant la probabilité pour qu'une mesure de A sur les deux particules fournisse la valeur a_n pour l'une et a_m pour l'autre. L'état issu de la mesure ayant fourni ces valeurs est lui aussi un état de deux particules identiques. En tant que tel, il doit satisfaire le postulat de symétrisation et est donc de la forme :

$$|\Psi'\rangle = \frac{1}{\sqrt{2}}\left[\mathbf{1} + \varepsilon\, T_{12}\right]|a_n a_m\rangle \ . \tag{22.67}$$

La probabilité de trouver ce résultat est donc :

$$P = |\langle\Psi'|\Psi\rangle|^2 = \left|\frac{1}{2}\,\langle a_n a_m|[\mathbf{1} + \varepsilon\, T_{12}^\dagger]^\dagger\,[\mathbf{1} + \varepsilon\, T_{12}]\,|\psi\phi\rangle\right|^2 \ ; \tag{22.68}$$

en vertu de $[\mathbf{1} + \varepsilon\, T_{12}]^\dagger = [\mathbf{1} + \varepsilon\, T_{12}^\dagger] = [\mathbf{1} + \varepsilon\, T_{12}]$, on obtient $[\mathbf{1} + \varepsilon\, T_{12}]^2$, qui vaut $2[\mathbf{1} + \varepsilon\, T_{12}]$ en raison de $T_{12}^2 = \mathbf{1}$; d'où :

$$P = |\langle a_n\, a_m|\psi\phi\rangle + \varepsilon\langle a_n\, a_m|\phi\,\psi\rangle|^2 \ . \tag{22.69}$$

Il apparaît ainsi deux termes : celui venant de $\mathbf{1}$ est dit *terme direct*, celui venant de T_{12} est appelé *terme d'échange*, pour des raisons évidentes. Ainsi, dans l'*amplitude* donnant la probabilité P, les bosons interfèrent avec le signe +, les fermions avec le signe − :

$$\boxed{P_{\text{bosons}} = |\langle a_n a_m|\psi\phi\rangle + \langle a_n a_m|\phi\,\psi\rangle|^2} \tag{22.70}$$

$$\boxed{P_{\text{fermions}} = |\langle a_n a_m|\psi\phi\rangle - \langle a_n a_m|\phi\,\psi\rangle|^2} \tag{22.71}$$

Ainsi, la probabilité la plus élémentaire de la Mécanique quantique affiche une expression qualitativement différente selon qu'il s'agit de fermions ou de bosons.

22.4 Etats d'un système de particules indépendantes : différence fondamentale entre bosons et fermions

On dit que N particules sont *indépendantes* s'il n'existe aucune interaction entre deux quelconques d'entre elles ; en conséquence, le Hamiltonien du système qu'elles constituent

ne contient aucun terme dépendant simultanément des grandeurs dynamiques (coordonnées, impulsions, spins, etc.) de deux particules (ou plus, dans le cas de modèles effectifs ou *renormalisés*). Il n'empêche que chacune d'entre elles peut être soumise individuellement à un champ de forces dérivant d'un potentiel et lui donnant l'énergie potentielle $V(\vec{r}_i)$. Dans ces conditions, H s'écrit :

$$H(1,2,\ldots,N) = \sum_{i=1}^{N} \left[\frac{\vec{p}_i^{\,2}}{2m} + V(\vec{r}_i)\right] \equiv \sum_{i=1}^{N} H_1(i) \ . \tag{22.72}$$

H est la somme de Hamiltoniens à une particule, H_1, tous identiques et ne différant dans la sommation que par l'indice i des variables dont ils dépendent. Désignons maintenant par ψ_k et E_k les modes propres de H_1 :

$$H_1\,\psi_k(\vec{x}) = E_k\psi_k(\vec{x}) \tag{22.73}$$

où $\vec{x} \equiv (\vec{r},\, m_s)$ dénote collectivement les coordonnées d'espace et la variable de spin. Si H_1 ne dépend pas du spin, la fonction ψ_k peut toujours être prise comme le produit d'une fonction ϕ_k par une fonction de spin χ, ϕ_k étant propre de H_1 :

$$\psi_k(\vec{x}) \equiv \phi_k(\vec{r}\,)\,\chi(m_s) \ , \qquad H_1\,\phi_k(\vec{r}\,) = E_k\,\phi_k(\vec{r}\,) \ . \tag{22.74}$$

Suivant la terminologie usuelle, la fonction $\psi_k(\vec{x})$ est appelée *spin-orbitale*, la fonction $\phi_k(\vec{r}\,)$ est appelée *orbitale*. Dans la suite, quand c'est possible, on continue à raisonner en toute généralité avec les spin-orbitales.

Comme H est une somme d'opérateurs H_1 à un corps, tout produit de spin-orbitales ψ_k sera fonction propre de H avec la valeur propre égale à la somme des valeurs propres associées aux spin-orbitales :

$$H\prod_{j=1}^{N}\psi_{k_j}(\vec{x}_j) = \sum_{i=1}^{N}\prod_{j=1}^{N} H_1(\vec{x}_i)\psi_{k_j}(\vec{x}_j) = \sum_{i=1}^{N}\Big[\prod_{j=1,\,j\neq i}^{N}\psi_{k_j}(\vec{x}_j)\Big]H_1(\vec{x}_i)\psi_{k_i}(\vec{x}_i) =$$
$$\sum_{i=1}^{N}\Big[\prod_{j=1,\,j\neq i}^{N}\psi_{k_j}(\vec{x}_j)\Big]E_{k_i}\psi_{k_i}(\vec{x}_i) = \sum_{i=1}^{N}E_{k_i}\prod_{j=1}^{N}\psi_{k_j}(\vec{x}_j) \ . \tag{22.75}$$

On a donc :

$$\boxed{H\prod_{j=1}^{N}\psi_{k_j}(\vec{x}_j) = E_{k_1,\,k_2,\,\ldots,\,k_N}\prod_{j=1}^{N}\psi_{k_j}(\vec{x}_j) \qquad E_{k_1,\,k_2,\,\ldots,\,k_N} = \sum_{j=1}^{N}E_{k_j}} \tag{22.76}$$

Cela étant, le produit $\prod_{j=1}^{N}\psi_{k_j}$ n'est pas une fonction physiquement acceptable vis-à-vis du postulat de symétrisation ; il convient maintenant de lui appliquer l'un des opérateurs $Y_{\rm I}$ (22.44) afin d'obtenir une fonction de symétrie convenable. Considérons d'abord le cas des bosons ; une bonne fonction est :

$$\Psi_{\rm S}(1,\,2,\,\ldots,\,N) = C_{\rm S}\,Y_{\rm S}\prod_{i=1}^{N}\psi_{k_i}(\vec{x}_i) \equiv C_{\rm S}\,Y_{\rm S}\,\Phi(1,\,2,\,\ldots,\,N) \tag{22.77}$$

l'énergie associée à Ψ_S étant la somme des E_{k_i}, conformément à (22.76). En particulier, l'état fondamental des N bosons sera construit en prenant pour chaque ψ_{k_i} l'état propre de H_1 de plus basse énergie, ψ_1, d'énergie E_1, donnant l'énergie totale $E_{1,1,\ldots,1} = N\, E_1$. Ainsi, l'état fondamental d'un système de bosons sans interactions est un état dans lequel tous les bosons sont chacun dans son état fondamental ; on dit que tous les bosons sont *condensés* dans le même état. Cette propriété remarquable ouvre la possibilité, à température finie et pour un système à nombre de bosons *fixé*[26], au phénomène appelé condensation de Bose[27]. L'état fondamental des N bosons est alors :

$$\Psi_{\text{fond bosons}} = \psi_1(\vec{x}_1)\,\psi_1(\vec{x}_2)\ldots\psi_1(\vec{x}_N) \ . \tag{22.78}$$

Cette fonction est visiblement symétrique et il ne servirait à rien de lui appliquer Y_S. Si les spin-orbitales $\psi_k(\vec{x})$ sont orthonormalisées, $\Psi_{\text{fond bosons}}$ est *de facto* normalisée à l'unité. Avec un choix quelconque des ψ_{k_i}, on obtient un état excité ; la fonction (22.77) est un *permanent*, somme de termes permutés sans aucun changement de signe (un permanent est comme un déterminant, mais sans l'alternance des signes).

Pour les fermions, il en va tout autrement. Comme la fonction d'onde totale doit être antisymétrique dans l'échange de deux fermions, on voit tout de suite que si deux fermions sont placés dans le même état quantique, la fonction d'onde est identiquement nulle. Pour bien voir ce fait fondamental, soit un système de deux fermions indépendants et plaçons-les tous deux dans le même état ψ_k (même partie orbitale, même valeur pour la projection du spin, plus généralement, même *spin-orbitale*). La fonction antisymétrique convenable s'obtient en appliquant Y_A qui, pour $N = 2$, s'écrit :

$$Y_A = \frac{1}{2!}\left[\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\right] , \tag{22.79}$$

d'où résulte :

$$\Psi_A(1,2) = C_A Y_A \psi_k(1)\psi_k(2) = \frac{1}{2}C_A\left[\psi_k(1)\psi_k(2) - \psi_k(2)\psi_k(1)\right] \equiv 0 \ . \tag{22.80}$$

Ainsi, pour des fermions, l'antisymétrie exige que deux d'entre eux n'occupent pas le même état quantique, faute de quoi l'état est identiquement nul – en d'autres termes

---

[26]Pas de condensation de Bose pour les photons, les phonons, etc.

[27]À température finie, le jeu des fluctuations thermiques peut supprimer la condensation de Bose si la dimensionnalité D est faible : l'argument *standard* strict – qui n'est pas exempt de critiques (Yvan Simon, communication privée) – montre qu'il n'y a pas de condensation de Bose si $D \leq 2$ (il donne $T_C \propto [\zeta(D/2)]^{-2/D}$ où ζ est la fonction de Riemann : pour $D \in \mathbb{R}$, $D \leq 2$, la fonction de Riemann est infinie et la température critique est *nulle*). Pour $D > 2$, la condensation se produit à une certaine température *finie* T_C : pour $T < T_C$, une fraction *macroscopique* de bosons est dans l'état fondamental ; cette fraction tend vers 1, quand la température tend vers zéro, suivant la loi-puissance $1 - (T/T_C)^{D/2}$.

Il est tout à fait remarquable qu'une transition de phase puisse se produire pour un système de particules sans interactions au sens classique. C'est une autre manifestation de l'"interaction" introduite par le 7e postulat, que l'on peut schématiquement qualifier d'attractive pour les bosons et de répulsive pour les fermions (donnant lieu au *trou de Fermi*, voir ci-dessous, p. 1013, et problème 22.6.1) ; d'ailleurs, quand on examine pour un gaz parfait presque classique (densité $\ll$ (longueur d'onde de de Broglie)$^{-D}$ $\propto T^{D/2}$) la première correction quantique à la fonction de partition, celle-ci peut s'interpréter en considérant une interaction effective (dépendant de la température...), attractive pour les bosons, répulsive pour les fermions (voir problème 22.6.6). Notons enfin que la question de la condensation de Bose en présence d'interactions entre les bosons (au sens ordinaire du terme) est encore un problème d'actualité.

n'existe pas en tant qu'état possible ; c'est ce qui fut historiquement (et empiriquement) énoncé sous la forme du *Principe d'exclusion de Pauli*, pour rendre compte des spectres atomiques, de la structure en couches des atomes et de la classification périodique.

Pour fabriquer l'état fondamental d'un système de fermions sans interactions, il faut donc prendre successivement tous les états propres de H_1 et placer dans chacun d'entre eux un et un seul fermion, compte tenu du spin. En notant ψ_1, ψ_2, ..., ψ_n, ... les états classés par ordre d'énergies croissantes $E_1 \leq E_2 \leq ... \leq E_n, ...$, l'état fondamental se construit à l'aide de Y_A et du simple produit $\psi_1\psi_2...\psi_N$: c'est bien l'état de plus basse énergie compatible avec le principe d'exclusion. Toutes choses égales par ailleurs[28], l'énergie totale d'un système de N fermions est donc considérablement plus grande que celle d'un système de N bosons puisque l'on est dans l'obligation d'aller chercher des états excités de H_1, chaque état propre de H_1 ne pouvant recevoir qu'un seul fermion. En tout cas, le fondamental s'écrit comme suit à l'aide des spin-orbitales :

$$\Psi_A(1, 2, ..., N) = C_A\, Y_A\, [\psi_1(\vec{x}_1)\psi_2(\vec{x}_2)...\psi_N(\vec{x}_N)] \equiv C_A Y_A \Phi(1,\, 2,\, ...,\, N)\ . \qquad (22.81)$$

Comme déjà mentionné, il arrive souvent que les états propres de H_1 ne dépendent pas du spin. En pareil cas, chaque ψ_k est le *produit* d'une orbitale $\phi_k(\vec{r})$ et d'une fonction de spin χ. Dans le cas d'électrons, $S = 1/2$, la spin-orbitale est soit $\phi_k(\vec{r})\alpha(m_s)$, soit $\phi_k(\vec{r})\beta(m_s)$, où les deux fonctions α et β ont été introduites dans le chapitre 20, éq. (20.36). On peut donc utiliser une fonction ϕ_k (orbitale) deux fois au plus, l'associant soit avec α, soit avec β. Ainsi, pour un nombre N pair d'électrons, tous ceux-ci sont appariés dans l'état fondamental et une bonne fonction pour cet état s'écrira :

$$\Psi_A(1, 2, ..., N) = C\, Y_A \prod_{i=1}^{\frac{N}{2}} \phi_i(\vec{r}_{2i-1})\alpha(m_{s\ 2i-1})\phi_i(\vec{r}_{2i})\beta(m_{s\ 2i})\ . \qquad (22.82)$$

Pour un nombre impair d'électrons, s'agissant toujours de l'état fondamental, le dernier état à une particule (celui de plus haute énergie) contiendra un seul électron non apparié, responsable d'un moment cinétique de spin total pour le système au moins égal[29] à $S = 1/2$. L'état occupé de plus haute énergie s'appelle le niveau de Fermi. Il joue un rôle prééminent dans un métal, comme on le verra au chapitre 30.

Ainsi, même s'ils sont sans interaction mutuelle au sens classique, deux fermions d'un même système sont malgré tout *corrélés* puisqu'ils ne peuvent se trouver dans le même état quantique. Par référence aux états orbitaux spécifiés par les trois nombres quantiques (n, l, m_l), auxquels il convient d'ajouter le nombre $m_s = \pm 1/2$ associé au spin, on dit que deux électrons d'un même atome, ou plus généralement d'un même système, ne peuvent avoir leurs quatre nombres quantiques identiques.

---

[28] notamment si la loi de dispersion est la même.

[29] Un sujet de réflexion : on admet généralement que dans la limite $N \ggg 1$, toutes les propriétés d'un système deviennent extensives en N et, notamment, sont *évidemment* (?) insensibles au fait que N est pair ou impair... Toutefois, quel que soit N, il existe ou il n'existe pas d'état singulet selon que N est pair ou impair, respectivement.

Revenons au cas de deux électrons, dont le spin total peut être 0 ou 1, suivant la composition des moments cinétiques ; avec deux fonctions d'espace ϕ_1 et ϕ_2 linéairement indépendantes, normalisées et orthogonales, on peut fabriquer un état singulet de spin total $S = 0$ et un état triplet de spin $S = 1$, ce dernier ayant trois composantes distinctes $M_S = 0, \pm 1$; la composante normalisée $M_S = +1$ du triplet est :

$$\Psi_{S=1,\, M_S=1}(1, 2) = \frac{1}{\sqrt{2}}\left[\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) - \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)\right]\alpha(m_{s_1})\alpha(m_{s_2}) \ , \tag{22.83}$$

et la probabilité de trouver les deux fermions, quelles que soient leurs projections de spin, l'un près de $\vec{r}_1$, l'autre près de $\vec{r}_2$ est :

$$\mathrm{d}P_{S=1}(\vec{r}_1, \vec{r}_2) = \frac{1}{2}\left|\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) - \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)\right|^2 \mathrm{d}^3r_1\mathrm{d}^3r_2 \ . \tag{22.84}$$

En particulier, la probabilité élémentaire est *nulle* de trouver les deux électrons au même point :

$$\boxed{\mathrm{d}P_{S=1}(\vec{r}_1 = \vec{r}_2) = 0} \tag{22.85}$$

C'est ce que l'on appelle le *trou de Fermi* : dans l'état triplet, il y a une probabilité nulle de trouver deux électrons au même point, ce que l'on résume souvent par une formule un peu trompeuse en disant que deux électrons de même spin ne peuvent se trouver au même point. Ce trou de densité n'existe pas pour l'état singulet, puisqu'alors la fonction d'onde est :

$$\Psi_{S=0, M_S=0}(1, 2) = \frac{1}{\sqrt{2}}\left[\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) + \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)\right]\frac{1}{\sqrt{2}}\left[\alpha(m_{s_1})\beta(m_{s_2}) - \beta(m_{s_1})\alpha(m_{s_2})\right] , \tag{22.86}$$

et la densité $\mathrm{d}P_{S=0}(\vec{r}_1, \vec{r}_2)$ ne s'annule pas en $\vec{r}_1 = \vec{r}_2$.

Ainsi, deux électrons "de même spin"[30] se repoussent davantage que deux électrons de spin contraire. Il est remarquable que cet antagonisme est indépendant de l'existence ou non d'une interaction ordinaire : il s'exerce en toute circonstance, et ne tient qu'au caractère fermionique ; il a une conséquence sur l'énergie du système dès qu'une interaction entre électrons est prise en compte (autrement, tous les états de spin sont dégénérés) – l'interaction la plus inévitable étant la répulsion électrostatique entre deux électrons. Comme celle-ci est essentiellement positive et est donc un facteur déstabilisant, on peut s'attendre à ce que, toutes choses égales par ailleurs, les états triplets soient plus bas en énergie que les états singulets. Ainsi, la seule interaction électrostatique, alliée au spin, assure que pour une configuration électronique donnée, l'état triplet est plus stable que l'état singulet : bien que le Hamiltonien ne contienne pas le spin, l'énergie totale dépend de S ! On retrouvera quantitativement ce résultat important au chapitre 27, dans le cadre le plus simple possible, celui de l'atome d'hélium (voir p. 1246).

C'est là une propriété fondamentale permettant de comprendre l'existence du paramagnétisme et plus généralement de mettre le doigt sur les briques élémentaires du magnétisme : les ampériens ne sont pas les fruits de forces magnétiques au sens

[30] Peut-on vraiment dire que dans la composante $M_S = 0$ du triplet les deux électrons ont le même spin ?

classique d'une interaction, ils résultent fondamentalement du caractère fermionique des électrons et existent même si ces derniers sont en interaction purement électrostatique, à l'exclusion de toute autre force de nature magnétique. Les corrélations dues à l'échange et à l'antisymétrisation, traduites par l'expression imagée *trou de Fermi* dans le cas des fermions, portent le nom de corrélations *dynamiques* ; on retiendra qu'elles sont présentes même en l'absence de toute interaction classique directe. En ce sens, qu'il s'agisse de bosons ou de fermions, des particules identiques sont toujours *forcément* corrélées.

À titre de transition vers le cas général de N particules, il est utile de remarquer que, pour $N = 2$, toute fonction antisymétrique peut s'écrire sous la forme d'un déterminant – qui est par définition une fonction (multilinéaire) antisymétrique. Ainsi, toute fonction de deux fermions construites sur deux spin-orbitales ψ_{k_1} et ψ_{k_2} (orthonormalisées) peut s'écrire :

$$\Psi_{\mathrm{A}}(1,2) = \frac{1}{\sqrt{2!}} \begin{vmatrix} \psi_{k_1}(1) & \psi_{k_2}(1) \\ \psi_{k_1}(2) & \psi_{k_2}(2) \end{vmatrix} \quad . \tag{22.87}$$

Il est facile de se convaincre que la fonction d'onde antisymétrique (22.81) construite sur un produit de N fonctions[31] à une particule $\{\psi_{k_i}\}$ peut de même s'écrire sous la forme d'un déterminant, historiquement appelé *déterminant de Slater* :

$$\Psi_{\mathrm{A}}(1,2,..., N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{k_1}(1) & \psi_{k_2}(1) & \ldots & \psi_{k_N}(1) \\ \psi_{k_1}(2) & \psi_{k_2}(2) & \ldots & \psi_{k_N}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{k_1}(N) & \psi_{k_2}(N) & \ldots & \psi_{k_N}(N) \end{vmatrix} \quad . \tag{22.88}$$

Bien évidemment, la permutation lignes $\leftrightarrow$ colonnes ne change rien. On vérifie aisément que $N!^{-1/2}$ est la bonne constante assurant que Ψ_{A} est normalisée à l'unité quand les spin-orbitales sont elles-mêmes orthonormalisées ; en effet, la condition de normalisation (22.59) avec $\Phi = \prod_{i=1}^{N} \psi_{k_i}$ s'écrit :

$$\langle \Psi_{\mathrm{A}} | \Psi_{\mathrm{A}} \rangle = \frac{1}{N!} |C_{\mathrm{A}}|^2 \sum_{\lambda=1}^{N!} \varepsilon_\lambda \langle \prod_{i=1}^{N} \psi_{k_i} | P_\lambda \prod_{i=1}^{N} \psi_{k_i} \rangle \quad . \tag{22.89}$$

La permutation P_λ transforme chaque indice i en i_λ ; ceci fait, les spin-orbitales du *bra* et du *ket* peuvent s'associer par paire selon leur ordre d'apparition, et le produit scalaire sur les N variables se factorise en N produits scalaires $\prod_{i=1}^{N} \langle \psi_{k_i} | \psi_{k_{i_\lambda}} \rangle$; grâce à l'orthogonalité des spin-orbitales, cette quantité est nulle sauf si $i_\lambda = i, \forall i$ – seul le terme venant de la permutation identité est non-nul –, d'où :

$$\langle \Psi_{\mathrm{A}} | \Psi_{\mathrm{A}} \rangle = |C_{\mathrm{A}}|^2 \frac{1}{N!} \prod_{i=1}^{N} \langle \psi_{k_i} | \psi_{k_i} \rangle = |C_{\mathrm{A}}|^2 \frac{1}{N!} = 1 \quad \Longleftrightarrow \quad |C_{\mathrm{A}}| = \sqrt{N!} \quad . \tag{22.90}$$

Finalement :

$$\boxed{|\Psi_{\mathrm{A}}\rangle = \frac{1}{\sqrt{N!}} \sum_{\lambda=1}^{N!} (-1)^{\sigma_\lambda} P_\lambda | \prod_{i=1}^{N} \psi_{k_i} \rangle} \tag{22.91}$$

[31] Bien noter que le déterminant est construit avec des spin-orbitales, pas avec des orbitales.

qui reproduit bien l'expression (22.88).

Sur la forme (22.88), on voit immédiatement qui si deux fonctions ψ_{k_i} coïncident (c'est-à-dire si deux *spin-orbitales* sont les mêmes), la fonction d'onde est identiquement nulle : alors, le déterminant a deux lignes (colonnes) identiques. La notation abrégée courante pour l'expression (22.88) est :

$$\Psi_\mathrm{A}(1, 2, ..., N) = \frac{1}{\sqrt{N!}}\,\mathrm{Det}\left|\psi_{k_1}\,\psi_{k_2}\,...\,\psi_{k_N}\right| \tag{22.92}$$

où apparaît explicitement la *diagonale* du déterminant, seule information à connaître pour reconstruire ce dernier.

Enfin, remarquons que tout opérateur qui est une somme d'opérateurs à une particule, $\Omega(1,\, 2,\, ...\, , N) \stackrel{\text{déf}}{=} \sum_{n=1}^{N} \Omega_1(n)$ a une action simple sur un déterminant ; puisqu'il s'agit d'une fonction multilinéaire, chaque terme de la somme n'agit que sur *une* ligne (ou une colonne, selon l'écriture adoptée) du déterminant, d'où :

$$\Omega(1,\, 2,\, ...\, , N)\,\mathrm{Det}\left|\psi_{k_1}\,\psi_{k_2}\,...\,\psi_{k_N}\right| = \sum_{n=1}^{N}\mathrm{Det}\left|\psi_{k_1}\,\psi_{k_2}\,...\,[\Omega_1(n)\psi_{k_n}],\, ...,\, \psi_{k_N}\right| \tag{22.93}$$

Si chaque spin-orbitale est propre de l'opérateur à une particule Ω_1 : $\Omega_1\psi_k = \omega_k\psi_k$, on obtient tout naturellement :

$$\Omega(1,\, 2,\, ...\, , N)\,\mathrm{Det}\left|\psi_{k_1}\,\psi_{k_2}\,...\,\psi_{k_N}\right| = \Big(\sum_{n=1}^{N}\omega_{k_n}\Big)\mathrm{Det}\left|\psi_{k_1}\,\psi_{k_2}\,...\,\psi_{k_N}\right| \;; \tag{22.94}$$

si Ω est le Hamiltonien, on retrouve le fait que l'antisymétrisation n'altère par la qualité d'état propre de la fonction Φ, l'énergie totale étant la somme des énergies individuelles, en conformité avec (22.76).

♦ *Remarques*

1. Le déterminant ci-dessus peut être construit à partir d'un modèle de particules indépendantes ; alors, chaque fonction ψ_{k_i} est connue d'avance. D'un autre côté, pour des fermions en interaction, on peut se poser la question de la meilleure fonction d'onde du type déterminant, au sens de la méthode variationnelle (voir chapitre 23). Dans une telle approche, les ψ_{k_i} sont au contraire inconnues *a priori* et doivent être déterminées. L'écriture du principe variationnel conduit alors à des équations intégro-différentielles pour les fonctions ψ_{k_i} : ce sont les célèbres équations de Hartree - Fock (voir sous-section 23.1.3), que l'on résout avec une machine, par un processus itératif. Cette méthode est typiquement une méthode de champ moyen : en effet, comme fondamentalement la fonction d'onde globale retient une forme *produit* de fonctions à une particule, chaque électron (fermion) se meut en réalité dans le champ *moyen* de tous les autres ; de fait, les équations de Hartree - Fock font apparaître pour les ψ_{k_i} une équation aux valeurs propres contenant un terme potentiel représentant l'effet moyen de tous les autres électrons

sur l'un d'entre eux[32]. En pareil cas, bien sûr, l'énergie des N électrons n'est pas la somme des valeurs propres (énergies) de Hartree - Fock – alors on compterait deux fois l'interaction au sein d'une même paire. La méthode de Hartree - Fock s'emploie pour toute assemblée de fermions (matière stellaire, noyaux, atomes, molécules, gaz de Fermi dans les solides, etc.)

2. Montrons que les fonctions antisymétriques obtenues pour deux électrons peuvent bien s'écrire sous la forme de déterminant(s). En effet, la fonction (22.83) est identique à $\frac{1}{\sqrt{2}}\text{Det}[\phi_1\alpha\,\phi_2\alpha]$. Comme le déterminant est une fonction multilinéaire (alternée), l'application de l'opérateur $S_- = S_{1-} + S_{2-}$ produit[33] (voir (22.93)) :

$$(S_{1-} + S_{2-})\,\frac{1}{\sqrt{2}}\,\text{Det}[\phi_1\alpha\,\phi_2\beta] = \frac{\hbar}{\sqrt{2}}\,[\text{Det}[\phi_1\beta\,\phi_2\alpha] + \text{Det}[\phi_1\alpha\,\phi_2\beta]\,] \quad . \qquad (22.95)$$

D'autre part :

$$S_-\Psi_{S=1\,M_S=1} = \hbar\sqrt{1(1+1) - 1(1-1)}\,\Psi_{S=1\,M_S=0} \equiv \hbar\sqrt{2}\,\Psi_{S=1\,M_S=0} \quad , \qquad (22.96)$$

de sorte que :

$$\Psi_{S=1\,M_S=0} = \frac{1}{2}\,[\text{Det}[\phi_1\beta\,\phi_2\alpha] + \text{Det}[\phi_1\alpha\,\phi_2\beta]\,] \quad . \qquad (22.97)$$

Il est facile de vérifier que ceci est identique à $\frac{1}{\sqrt{2}}(\phi_1\phi_2 - \phi_2\phi_1)\,\frac{1}{\sqrt{2}}(\alpha\beta + \beta\alpha)$; la composante $M_S = -1$ du triplet est évidemment $\frac{1}{\sqrt{2}}\text{Det}[\phi_1\beta\,\phi_2\beta]$. Pour obtenir la seule et unique composante du singulet $\Psi_{S=0\,M_S=0}$, il suffit de se souvenir que cette fonction doit être orthogonale à $\Psi_{S=1\,M_S=0}$. Deux déterminants différant par au moins une spin-orbitale sont orthogonaux si les spin-orbitales le sont ; la fonction orthogonale à (22.97) est donc :

$$\Psi_{S=0\,M_S=0} = \frac{1}{2}\,[\text{Det}[\phi_1\beta\,\phi_2\alpha] - \text{Det}[\phi_1\alpha\,\phi_2\beta]] \quad , \qquad (22.98)$$

obtenue par un simple changement de signe dans la combinaison linéaire des deux déterminants ; cette fonction est bien égale à $\frac{1}{\sqrt{2}}(\phi_1\phi_2 + \phi_2\phi_1)\,\frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha)$.

On retrouve ainsi que, pour $N = 2$, les parties d'espace et de spin restent *factorisées* après antisymétrisation (elles ne sont pas *intriquées*), une propriété qui disparaît si $N > 2$, sauf pour l'état *ferromagnétique*, correspondant à la valeur maximale du spin, soit $S = N/2$. ♦

[32] De façon *caractéristique* pour un système de fermions, le potentiel de champ moyen fait apparaître *deux* termes ; le premier (terme *direct*) a une interprétation simple : c'est l'interaction d'un électron avec les autres, distribués suivant une densité donnée par la somme des modules carrés des spin-orbitales. En revanche, le second (terme d'*échange*) résulte directement du postulat d'antisymétrisation et ne peut recevoir d'interprétation imagée. L'approximation dite de Hartree consiste à oublier le terme d'échange.

[33] étant entendu que $S_-\alpha = \hbar\sqrt{\frac{1}{2}(\frac{1}{2}+1) - \frac{1}{2}(\frac{1}{2}-1)}\,\beta \equiv \hbar\beta$.

22.5 Introduction à la Seconde quantification

Compte tenu de l'indiscernabilité des particules identiques, la spécification complète de l'état d'un système s'énonce en affirmant que n_1 particules sont dans l'état ψ_1, n_2 particules sont dans l'état ψ_2, etc.. Souvent, les états à une particule ψ_k sont ceux définis en l'absence d'interaction, mais ceci n'est nullement obligatoire : en présence d'interactions, on redéfinit des *quasi-particules*, qui peuvent de fait avoir des propriétés très différentes des composés élémentaires qui les constituent (deux fermions donnent un boson) ; une fois ces quasi-particles bien définies avec leurs états propres, l'état du système se décrit suivant la même procédure, en donnant les nombres d'occupation de ces excitations *rhabillées* par les interactions des constituants élémentaires qui les composent. Dans la suite, l'indice k désigne ainsi généralement l'ensemble des nombres quantiques associés à un ECOC spécifiant complètement l'état (*mode*) d'une particule, ou d'une quasi-particule.

Par ailleurs, on est souvent contraint en pratique de considérer des situations où le nombre de particules n'est pas fixé une fois pour toutes. Par exemple, dans le cas du champ électromagnétique couplé à la matière, le nombre de photons n'est pas donné et constant puisque la matière peut absorber ou émettre sans cesse des photons, que ce soit de façon réelle ou virtuelle. Il en va de même pour les phonons, particules (*quanta*) décrivant la quantification des vibrations harmoniques d'un solide ; le nombre (moyen) de phonons, qui augmente avec la température puisque l'élévation de celle-ci augmente l'amplitude des vibrations du réseau (voir chapitre 31), présente inévitablement des *fluctuations* ($\langle n^2\rangle \geq \langle n\rangle^2$). En Théorie quantique des champs, la non-conservation du nombre de particules au cours du temps est la règle : une paire électron-positron peut s'annihiler sous forme de photons, le processus inverse étant tout autant possible. C'est même plus qu'une règle : toute cette théorie est construite sur le concept d'un état fondamental (vide) et de ses excitations ; par *construction*, le formalisme incorpore d'emblée *tous* les états à nombre quelconque de particules. Dès lors, une particule apparaît comme une *excitation élémentaire* du champ considéré.

Ainsi survient la nécessité de reformuler la Mécanique quantique dans un cadre plus large où le nombre de particules est variable[34] et dans lequel l'espace d'états est l'*espace de Fock*, déjà introduit (note 77, p. 827). On peut se le représenter comme un empilement de feuilles[35] (secteurs) : dans chaque feuille, le nombre de particules est constant ; on passe d'une feuille à l'autre par l'action d'opérateurs de création ou d'annihilation de particules, qui généralisent ceux que l'on a rencontrés à propos de l'oscillateur harmonique, et dont il existe évidemment deux familles, l'une associée aux bosons, l'autre associée aux fermions[36]. On verra que la distinction se tient fondamentalement dans les relations de commutation : les opérateurs de bosons satisfont des règles de commutation (exactement comme les a et $a^\dagger$ de l'oscillateur harmonique), les opérateurs de fermions obéissent à des relations d'*anticommutation*. Cette distinction contient à elle

[34]Pour en savoir plus sur la Seconde quantification, voir l'ouvrage de Fetter et Walecka [174].

[35]Une image : l'espace de Fock ressemble à la surface de Riemann de la fonction multiforme $\ln z$ quand on s'autorise à tourner dans le sens positif exclusivement ; un secteur de Fock correspond à un feuillet de Riemann, où l'angle est choisi dans l'intervalle $[2n\pi,\ (2n+2)\pi]$, l'entier $n \in \mathbb{N}$ étant fixé.

[36]Tout *anyon* mis à part...

seule la différence de symétrie vis-à-vis du groupe de permutations, en conformité avec le postulat de symétrisation.

Les états (modes propres) de chaque particule étant définis (nus ou renormalisés), un état quelconque s'obtient en effectuant le produit tensoriel des états contenant chacun un nombre donné de particules. En omettant la notation $\otimes$ pour simplifier l'écriture, un état quelconque défini dans l'espace de Fock peut être noté :

$$|n_1, n_2, ..., n_k, ...\rangle \tag{22.99}$$

où les n_i sont des entiers positifs ou nuls. Pour les bosons, l'ordre d'apparition des entiers non-nuls dans le *ket* est sans importance, puisque les deux états $|..., n_k, ..., n_l, ...\rangle$ et $|..., n_l, ..., n_k, ...\rangle$ sont égaux en raison de la symétrie de permutation. Au contraire, pour des fermions, les deux états $|..., 1_k, ..., 1_l, ...\rangle$ et $|..., 1_l, ..., 1_k, ...\rangle$ diffèrent d'un signe en vertu de l'antisymétrisation. En fonction du contexte, il pourra être nécessaire d'utiliser une certaine convention d'écriture conduisant à une notation non-ambiguë d'un état de fermions ; la plus usuelle, et la plus naturelle, consiste à ranger les n_k de gauche à droite par ordre d'énergies croissantes des spin-orbitales correspondantes, convention complétée par une autre en cas de dégénérescence.

On introduit ensuite des opérateurs de création $a_k^\dagger$ et d'annihilation a_k tels que :

$$a_k^\dagger|n_1, n_2, ..., n_k, ...\rangle \propto |n_1, n_2, ..., n_{k-1}, n_k + 1, n_{k+1}, ...\rangle \ , \tag{22.100}$$

$$a_k|n_1, n_2, ..., n_k, ...\rangle \propto |n_1, n_2, ..., n_{k-1}, n_k - 1, n_{k+1}, ...\rangle \ . \tag{22.101}$$

Ce sont ces opérateurs qui permettent de passer d'un secteur à l'autre de l'espace de Fock. Chaque opérateur a_k et $a_k^\dagger$ n'agit que sur son mode k et ne modifie que le nombre d'occupation n_k, le faisant varier exactement d'une unité, dans un sens ou dans l'autre – sans doute la traduction technique la plus simple de la notion de *quantum*.

Cela étant posé, il est possible de distinguer notamment :

L'état vide

L'état vide, ne contenant aucune excitation (particule), décrit l'état fondamental du champ quantique associé aux opérateurs $\{a_k\}_k$ et $\{a_k^\dagger\}_k$:

$$|\text{vide}\rangle \equiv |n_1 = 0, n_2 = 0, ..., n_k = 0, ...\rangle \ . \tag{22.102}$$

Pour l'état vide, on a :

$$\boxed{a_k|\text{vide}\rangle = 0} \tag{22.103}$$

Cette relation exprime le fait que l'on ne peut détruire une particule qui n'existe pas. Il n'existe pas d'état d'énergie inférieure à celle du vide, laquelle est usuellement prise comme zéro d'énergie.

Les états à une particule

Ce sont les états où tous les n_k sont nuls, sauf un, égal à 1 :

$$|\Psi_k\rangle \equiv |n_1 = 0,\, n_2 = 0,\, ...n_{k-1} = 0,\, n_k = 1,\, n_{k+1} = 0,\, ...\rangle \ , \qquad (22.104)$$

que l'on peut construire en faisant agir une fois un opérateur de création sur l'état vide, conformément à (22.100) :

$$\boxed{|\Psi_k\rangle = a_k^\dagger|\text{vide}\rangle} \qquad (22.105)$$

Par ailleurs, toujours grâce à l'argument suivant lequel il est impossible de détruire une particule qui n'existe pas, on a :

$$a_k|0,\, 0,\, ...,\, n_{k'},\, 0,\, ...\rangle = 0 \qquad \forall\, k \neq k' \ . \qquad (22.106)$$

Les états à deux particules

Les états à deux particules, $|n_k = 1,\, n_l = 1\rangle$ sont engendrés à partir du vide par action d'un produit tel que $a_k^\dagger a_l^\dagger$:

$$\boxed{|\Psi_{kl}\rangle = a_k^\dagger a_l^\dagger|\text{vide}\rangle} \qquad (22.107)$$

Si les particules sont des bosons, l'état doit être symétrique ; l'égalité suivante doit donc être vraie :

$$a_k^\dagger a_l^\dagger|\text{vide}\rangle = a_l^\dagger a_k^\dagger|\text{vide}\rangle \quad \forall\, k,\, l \ . \qquad (22.108)$$

Elle doit l'être d'ailleurs quel que soit l'état sur lequel agit le produit $a_k^\dagger a_l^\dagger$, puisqu'un état de bosons, quel qu'il soit, doit être symétrique dans toute transposition. On en déduit, pour des bosons :

$$a_k^\dagger a_l^\dagger = a_l^\dagger a_k^\dagger \quad \forall\, k,\, l \quad \Longleftrightarrow \quad \boxed{[a_k^\dagger,\, a_l^\dagger] = 0 \quad (\text{bosons})} \qquad (22.109)$$

La relation hermitique conjuguée est tout aussi vraie :

$$a_k a_l = a_l a_k \quad \forall\, k,\, l \quad \Longleftrightarrow \quad \boxed{[a_k,\, a_l] = 0 \quad (\text{bosons})} \qquad (22.110)$$

Au contraire, s'il s'agit de fermions, l'état doit être antisymétrique dans toute transposition ; il faut que :

$$a_k^\dagger a_l^\dagger|\text{vide}\rangle = -\, a_l^\dagger a_k^\dagger|\text{vide}\rangle \quad \forall\, k,\, l \qquad (22.111)$$

et, plus généralement[37] :

$$a_k^\dagger a_l^\dagger = -\, a_l^\dagger a_k^\dagger \quad \forall\, k,\, l \quad \Longleftrightarrow \quad \boxed{\{a_k^\dagger,\, a_l^\dagger\} = 0 \quad (\text{fermions})} \qquad (22.112)$$

[37] Les accolades désignent l'anticommutateur : $\{A,\, B\} \overset{\text{déf}}{=} AB + BA$.

qui entraîne à son tour par conjugaison hermitique :

$$a_k a_l = -\,a_l a_k \quad \forall\, k,\, l \quad \Longleftrightarrow \quad \boxed{\{a_k,\, a_l\} = 0 \quad \text{(fermions)}} \tag{22.113}$$

Ainsi, le produit d'un nombre $\mathcal{N}$ quelconque d'opérateurs de création agissant sur le vide permet de construire les états à $\mathcal{N}$ particules. L'ensemble de tous ces états constitue l'espace de Fock, qui est de toute évidence un espace de Hilbert, espace vectoriel sur $\mathbb{C}$ muni d'un produit scalaire sesquilinéaire et physiquement supposé complet.

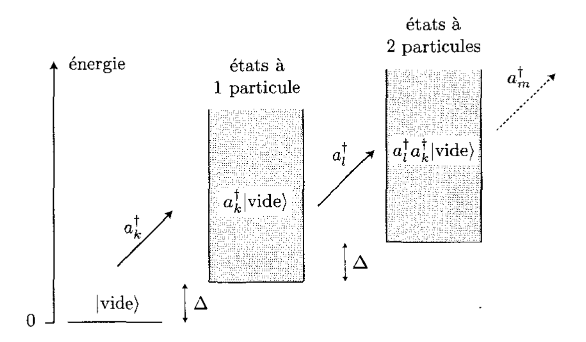


Figure 22.3: Représentation schématique du spectre des différents états de l'espace de Fock, dans le cas d'excitations de masse finie (avec *gap*) dont la création exige l'énergie au moins égale à Δ.

Il convient maintenant de déterminer les constantes de proportionnalité dans (22.100) et (22.101), relations que l'on peut écrire plus précisément (compte tenu du fait que a_k et $a_k^\dagger$ sont hermitiques conjugués l'un de l'autre) :

$$a_k^\dagger|...,\, n_k, ...\rangle \;=\; C_{n_k+1}\,|...,n_k+1,\, ...\rangle \;, \tag{22.114}$$

$$a_k|...,\, n_k, ...\rangle \;=\; C_{n_k}^*\,|...,\, n_k-1,\, ...\rangle \;. \tag{22.115}$$

Tous les nombres $n_{k'\neq k}$ sont sous-entendus dans les *kets*, et inchangés d'un membre à l'autre ; les deux constantes ne dépendent évidemment que du nombre d'occupation du mode que les opérateurs modifient (deux modes distincts ne se voient pas). De ces deux relations on déduit :

$$a_k^\dagger a_k|...,\, n_k, ...\rangle = C_{n_k}^*\, a_k^\dagger|...,\, n_k-1, ...\rangle = |C_{n_k}|^2\,|...,\, n_k, ...\rangle \;. \tag{22.116}$$

Ainsi, les états de Fock sont propres de $N_k \stackrel{\text{déf}}{=} a_k^\dagger a_k$. L'opérateur N_k généralise pour le mode k l'opérateur $a^\dagger a$ introduit à propos de la description de Dirac d'un oscillateur harmonique à une dimension (Tome I, section 16.3). On peut alors développer, pour le mode k, exactement le même argument que pour l'oscillateur harmonique et arriver à la conclusion que l'opérateur hermitique N_k a pour valeurs propres tous les entiers positifs ou nuls, d'où $|C_{n_k}|^2 = n_k \in \mathbb{N}$. Pour cette raison, N_k est appelé *opérateur nombre*

d'occupation (de l'état k) ; sa valeur propre donne le nombre de particules (*quanta*) effectivement placées dans l'état k :

$$\boxed{N_k = a_k^\dagger a_k \qquad N_k|..., n_k, ...\rangle = n_k\,|..., n_k, ...\rangle} \tag{22.117}$$

Les relations (22.114), (22.115) et (22.116) sont vraies qu'il s'agisse de fermions ou de bosons, mais pour la clarté, traitons maintenant séparément chaque famille de particules.

Bosons

Pour les bosons, chaque état k peut accepter un nombre quelconque de particules ; la relation $n_k = |C_{n_k}|^2$ permet alors de fixer (à une phase près conventionnelle) la constante C_{n_k}. En décidant de prendre cette dernière réelle positive, il vient $C_{n_k} = \sqrt{n_k}$; selon (22.114) et (22.115), il vient :

$$\boxed{a_k^\dagger|..., n_k, ...\rangle = \sqrt{n_k+1}\,|..., n_k+1, ...\rangle} \tag{22.118}$$

$$\boxed{a_k|..., n_k, ...\rangle = \sqrt{n_k}\,|..., n_k-1, ...\rangle} \tag{22.119}$$

La combinaison de ces deux équations redonne bien (22.116) :

$$a_k^\dagger a_k|..., n_k, ...\rangle = \sqrt{n_k}\,a_k^\dagger|..., n_k-1, ...\rangle = \sqrt{n_k}\,\sqrt{1+(n_k-1)}|..., (n_k-1)+1, ...\rangle \tag{22.120}$$

Ceci étant établi, on voit que, si $l \neq k$:

$$a_l a_k^\dagger|..., n_k, ..., n_l, ...\rangle = \sqrt{n_l}\sqrt{n_k+1}|..., n_k+1, ..., n_l-1, ...\rangle \ ; \tag{22.121}$$

le produit renversé, $a_k^\dagger a_l$, donne manifestement le même résultat, d'où $[a_k, a_l^\dagger] = 0$ si $l \neq k$. Au contraire, pour $l = k$, il vient :

$$a_k a_k^\dagger|..., n_k, ...\rangle = \sqrt{n_k+1}\sqrt{n_k+1}|..., n_k, ...\rangle \ ; \tag{22.122}$$

cependant que :

$$a_k^\dagger a_k|..., n_k, ...\rangle = \sqrt{(n_k-1)+1}\sqrt{n_k}|..., n_k, ...\rangle \ , \tag{22.123}$$

la différence membre à membre donne simplement $|..., n_k, ...\rangle$ au second membre, d'où, au total :

$$\boxed{[a_k, a_l^\dagger\,] \ = \delta_{kl}\mathbf{1} \qquad \text{(bosons)}} \tag{22.124}$$

Ainsi, pour les bosons, on retrouve pour chaque mode k toutes les relations caractéristiques rencontrées avec l'oscillateur harmonique (Tome I, section 16.3), qui constitue en fait à lui tout seul *un* certain mode bosonique.

Fermions

Pour les fermions, il reste vrai que $|C_{n_k}|^2$ est une fonction linéaire de n_k, en accord avec (22.116), mais maintenant n_k ne peut prendre que les deux valeurs 0 ou 1, en conformité avec le Principe de Pauli, ce qui conduit à des relations *d'anticommutation.* Partant de l'état vide, on peut toujours définir les phases de sorte que :

$$a_k^\dagger|\text{vide}\rangle = |1_k\rangle \quad \Longleftrightarrow \quad a_k|1_k\rangle = |\text{vide}\rangle \ . \tag{22.125}$$

Examinons maintenant les produits $a_k a_k^\dagger$ et $a_k^\dagger a_k$; agissant sur le vide et selon (22.125), ils donnent :

$$a_k a_k^\dagger|\text{vide}\rangle = a_k|1_k\rangle = |\text{vide}\rangle \ , \qquad a_k^\dagger a_k|\text{vide}\rangle = 0 \ , \tag{22.126}$$

d'où $(a_k a_k^\dagger \pm a_k^\dagger a_k)|\text{vide}\rangle = |\text{vide}\rangle$. Calculons maintenant l'action de ces produits sur l'état à une particule $|1_k\rangle$:

$$a_k a_k^\dagger|1_k\rangle = 0 \ , \qquad a_k^\dagger a_k|1_k\rangle = a_k^\dagger|\text{vide}\rangle = |1_k\rangle \ , \tag{22.127}$$

d'où $(a_k a_k^\dagger + a_k^\dagger a_k)|1_k\rangle = |1_k\rangle$. Maintenant, pour un état quelconque $|n_k,\, \{n_l\}_{l\neq k}\rangle$, ces mêmes opérateurs n'affectent que le mode k :

$$a_k^\dagger|n_k,\, \{n_l\}_{l\neq k}\rangle\rangle = \sqrt{1-n_k}\,|n_k+1,\, \{n_l\}_{l\neq k}\rangle \ , \tag{22.128}$$

$$a_k|n_k,\, \{n_l\}_{l\neq k}\rangle\rangle = \sqrt{n_k}\,|n_k-1,\, \{n_l\}_{l\neq k}\rangle \ , \tag{22.129}$$

et on obtient exactement les mêmes résultats, la distribution $\{n_l\}_{l\neq k}$ apparaissant inaltérée partout dans les *kets*. On en déduit, au niveau des opérateurs eux-mêmes, l'égalité $\{a_k,\, a_k^\dagger\} = \mathbf{1}$.

Soit maintenant les produits $a_k a_l^\dagger$ et $a_l^\dagger a_k$, $l \neq k$, que l'on va faire agir sur des états à nombre donné de fermions. Sur le vide, on a :

$$a_k a_l^\dagger|\text{vide}\rangle = a_k|1_l\rangle = 0 \ , \qquad a_l^\dagger a_k|\text{vide}\rangle = 0 \ , \tag{22.130}$$

d'où $(a_k a_l^\dagger \pm a_l^\dagger a_k)|\text{vide}\rangle = 0$. Sur l'état $|1_m\rangle$, ils donnent[38] :

$$a_k a_l^\dagger|1_m\rangle = (1-\delta_{lm})a_k|1_l 1_m\rangle = (1-\delta_{lm})(\delta_{kl}|1_m\rangle - \delta_{km}|1_l\rangle) = -(1-\delta_{lm})\delta_{km}|1_l\rangle \ , \tag{22.131}$$

la dernière égalité venant de l'hypothèse $k \neq l$. Par ailleurs, $a_l^\dagger a_k|1_m\rangle = \delta_{km}|1_l\rangle$, d'où :

$$(a_k a_l^\dagger + a_l^\dagger a_k)|1_m\rangle = \delta_{km}\delta_{lm}|1_l\rangle \ ; \tag{22.132}$$

comme $k \neq l$, le produit des deux symboles de Kronecker est toujours nul.

Faisons agir ces produits sur un état à deux fermions, $|1_{k'}1_{l'}\rangle$; on a :

$$a_k a_l^\dagger|1_{k'}1_{l'}\rangle = a_k[(1-\delta_{lk'}) + (1-\delta_{ll'})]|1_l 1_{k'} 1_{l'}\rangle \ ; \tag{22.133}$$

[38] Se souvenir de $a_k|1_k 1_l\rangle = |1_l\rangle$ et de $a_l|1_k 1_l\rangle = -a_l|1_l 1_k\rangle = -|1_k\rangle$.

on a $a_k|1_l 1_{k'} 1_{l'}\rangle = \delta_{kl}|1_{k'}1_{l'}\rangle - \delta_{kk'}|1_l 1_{l'}\rangle + \delta_{kl'}|1_l 1_{k'}\rangle = -\delta_{kk'}|1_l 1_{l'}\rangle + \delta_{kl'}|1_l 1_{k'}\rangle$. Par ailleurs :

$$a_l^\dagger a_k|1_{k'}1_{l'}\rangle = a_l^\dagger(\delta_{kk'}|1_{l'}\rangle - \delta_{kl'}|1_{k'}\rangle) = \delta_{kk'}(1-\delta_{ll'})|1_l 1_{l'}\rangle - \delta_{kl'}(1-\delta_{lk'})|1_l 1_{k'}\rangle \;\; ; \quad (22.134)$$

faisant la somme, on trouve que $a_k a_l^\dagger + a_l^\dagger a_k$ agissant sur un état quelconque à deux particules donne zéro, si $k \neq l$, et ainsi de suite : l'application à n'importe quel autre état donne toujours le même résultat. Au total, on a donc :

$$\boxed{\{a_k,\, a_l^\dagger\} = \delta_{kl}\mathbf{1} \qquad \text{(fermions)}} \quad (22.135)$$

$N_k \overset{\text{déf}}{=} a_k^\dagger a_k$ est dans tous les cas (bosons et fermions) l'opérateur nombre de particules dans le mode k ; ses valeurs propres sont des entiers non-négatifs :

$$\boxed{N_k = \begin{cases} 0,\,1,\,2,\,... & \text{(bosons)} \\ 0,\,1 & \text{(fermions)} \end{cases}} \quad (22.136)$$

L'usage courant désigne par la lettre b les opérateurs de bosons, par c ceux des fermions. Avec cette convention, et à titre de récapitulation, les relations fondamentales s'écrivent :

Bosons

$$\forall k,\, l \,:\quad [b_k,\, b_l^\dagger] = \delta_{kl}\mathbf{1} \;\;, \quad [b_k,\, b_l] = [b_k^\dagger,\, b_l^\dagger] = 0 \;\;, \quad (22.137)$$

$$b_k^\dagger\,|n_1,\, n_2,\, ...,\, n_k,\, n_{k+1},\, ...\rangle = \sqrt{n_k+1}\,|n_1,\, n_2,\, ...,\, n_k+1,\, n_{k+1},\, ...\rangle \;\;. \quad (22.138)$$

$$b_k\,|n_1,\, n_2,\, ...,\, n_k,\, n_{k+1},\, ...\rangle = \sqrt{n_k}\,|n_1,\, n_2,\, ...,\, n_k-1,\, n_{k+1},\, ...\rangle \;\;, \quad (22.139)$$

Fermions

$$\forall k,\, l \,:\quad \{c_k,\, c_l^\dagger\} = \delta_{kl}\mathbf{1} \;\;, \quad \{c_k,\, c_l\} = \{c_k^\dagger,\, c_l^\dagger\} = 0 \;\;, \quad (22.140)$$

$$c_k^\dagger\,|n_1,\, n_2,\, ...,\, n_k,\, n_{k+1},\, ...\rangle = (-1)^{\alpha_k}\sqrt{1-n_k}\,|n_1,\, n_2,\, ...,\, n_k+1,\, n_{k+1},\, ...\rangle \quad (22.141)$$

$$c_k\,|n_1,\, n_2,\, ...,\, n_k,\, n_{k+1},\, ...\rangle = (-1)^{\alpha_k}\sqrt{n_k}\,|n_1,\, n_2,\, ...,\, n_k-1,\, n_{k+1},\, ...\rangle \;\;, \quad (22.142)$$

Afin d'expliquer l'origine de la phase $(-1)^{\alpha_k}$ – liée à la convention d'ordre sur les états à une particule –, revenons à (22.128) et récrivons cette équation dans les deux cas possibles :

$$n_k = 0 \,:\, c_k^\dagger|0_k,\, \{n_l\}_{l\neq k}\rangle\rangle = |1_k,\, \{n_l\}_{l\neq k}\rangle \;\;, \quad n_k = 1 \,:\, c_k^\dagger|1_k,\, \{n_l\}_{l\neq k}\rangle\rangle = 0 \quad (22.143)$$

Adoptons un ordre conventionnel pour les états (par exemple, celui des énergies croissantes, $E_k < E_{k+1}$, $k \in \mathbb{N}^*$) ; la notation *naturelle* des états de Fock est $|n_1, n_2, ..., n_k, ...\rangle$. Dans le cas $n_k = 1$, la question de retrouver l'ordre adopté est sans intérêt puisque le second membre est nul. Quand $n_k = 0$, notons d'abord que le réordonnancement de l'ensemble $\{n_l\}_{l\neq k}$ donne la même phase aux deux membres : on peut donc l'oublier. Quant au *ket* au second membre, il retrouve la notation ordonnée en faisant faire à l'indice

k des sauts de puce afin de remettre celui-ci à la bonne place, chaque saut correspondant à une certaine transposition ; ces transpositions ne sauraient impliquer un état qui est vide dans la collection $\{n_l\}_{l\neq k}$: les seules transpositions à effectuer sont celles pour lesquelles $n_l = 1$, avec $l < k$, de sorte que la somme $\alpha_k \stackrel{\text{déf}}{=} n_1 + n_2 + \ldots + n_{k-1}$ sert de fonction indicatrice du nombre de transpositions à effectuer. Chaque transposition donnant un facteur (-1), l'écriture avec la notation conventionnelle fait apparaître la phase $(-1)^{\alpha_k}$ en facteur du *ket* au second membre. Pour l'autre égalité (22.129), c'est cette fois le cas $n_k = 1$ qui importe :

$$c_k|1_k,\,\{n_l\}_{l\neq k}\rangle\rangle = |0_k,\,\{n_l\}_{l\neq k}\rangle \ ; \tag{22.144}$$

on retrouve le même *ket* à réordonner – au premier membre –, d'où à nouveau la phase $(-1)^{\alpha_k}$ dans (22.142).

Le point essentiel de l'écriture en Seconde quantification est que la symétrie de permutation est *incluse* dans l'algèbre des opérateurs de création et d'annihilation : commutation pour les bosons, anticommutation pour les fermions. De surcroît, la notion de *champ* émerge maintenant de façon très naturelle ; afin de montrer ceci sans complications d'écriture, prenons le cas de bosons de spin nul[39]. Dans ces conditions, les états d'un boson sont les simples orbitales $\psi_k(\vec{r})$, le nombre quantique k variant dans un certain ensemble K, et ces fonctions satisfaisant la relation de fermeture en représentation-q :

$$\sum_{k\in K} \psi_k(\vec{r})\psi_k^*(\vec{r}\,') = \delta(\vec{r}-\vec{r}\,') \ . \tag{22.147}$$

Définissons maintenant les *opérateurs de champ* :

$$\hat{\psi}(\vec{r}) \stackrel{\text{déf}}{=} \sum_{k\in K} \psi_k(\vec{r})\, b_k \ , \qquad \hat{\psi}^\dagger(\vec{r}) \stackrel{\text{déf}}{=} \sum_{k\in K} \psi_k^*(\vec{r})\, b_k^\dagger \ ; \tag{22.148}$$

de ces définitions, et en vertu de (22.137) et (22.147), résulte immédiatement :

$$\boxed{[\hat{\psi}(\vec{r}),\, \hat{\psi}(\vec{r}\,')] = [\hat{\psi}^\dagger(\vec{r}),\, \hat{\psi}^\dagger(\vec{r}\,')] = 0 \qquad [\hat{\psi}(\vec{r}),\, \hat{\psi}^\dagger(\vec{r}\,')] = \delta(\vec{r}-\vec{r}\,')\mathbf{1}} \tag{22.149}$$

Compte tenu de ceci, on a :

$$\langle \text{vac}|\hat{\psi}(\vec{r})\hat{\psi}^\dagger(\vec{r}\,')|\text{vac}\rangle = \delta(\vec{r}-\vec{r}\,') \tag{22.150}$$

[39]D'une façon générale, pour des particules de spin S quelconque, on définit un opérateur de champ pour chaque composante M_S du multiplet de spin ($-S \leq M_S \leq +S$). Notant simplement σ la valeur de M_S, on définit :

$$\hat{\psi}_\sigma(\vec{r}) \stackrel{\text{déf}}{=} \sum_{k\in K} \psi_{k\sigma}(\vec{r})\, b_{k\sigma} \ , \qquad \hat{\psi}_\sigma^\dagger(\vec{r}) \stackrel{\text{déf}}{=} \sum_{k\in K} \psi_{k\sigma}^*(\vec{r})\, b_{k\sigma}^\dagger \ ; \tag{22.145}$$

avec $[b_{k\sigma},\, b_{k'\sigma'}]_\mp = [b_{k\sigma}^\dagger,\, b_{k'\sigma'}^\dagger]_\mp = 0$, $[b_{k\sigma},\, b_{k'\sigma'}^\dagger]_\mp = \delta_{kk'}\delta_{\sigma\sigma'}\mathbf{1}$; cela étant, les relations de commutation des opérateurs de champ sont :

$$[\hat{\psi}_\sigma(\vec{r}),\, \hat{\psi}_{\sigma'}(\vec{r}\,')]_\mp = [\hat{\psi}_\sigma^\dagger(\vec{r}),\, \hat{\psi}_{\sigma'}^\dagger(\vec{r}\,')]_\mp = 0 \qquad [\hat{\psi}_\sigma(\vec{r}),\, \hat{\psi}_{\sigma'}^\dagger(\vec{r}\,')]_\mp = \delta(\vec{r}-\vec{r}\,')\delta_{\sigma\sigma'}\,\mathbf{1} \tag{22.146}$$

où $[A,\, B]_\mp \stackrel{\text{déf}}{=} AB \mp BA$ et où le signe supérieur correspond aux bosons, le signe inférieur aux fermions.

ce qui permet d'interpréter $\hat{\psi}^\dagger(\vec{r})$ comme l'opérateur créant une particule au point $\vec{r}$; la localisation extrême vient du fait que $\hat{\psi}^\dagger(\vec{r})$ contient *tous* les états d'une particule, supposés complets.

Un autre intérêt de la Seconde quantification est que les observables s'y expriment très simplement, et de façon très transparente. Soit $|\Psi_k\rangle$ un état à une particule :

$$|\Psi_k\rangle = a_k^\dagger |\text{vac}\rangle \ , \tag{22.151}$$

propre du Hamiltonien H_1 :

$$H_1|\Psi_k\rangle = \varepsilon_k\,|\Psi_k\rangle \quad \Longleftrightarrow \quad H_1 = \sum_k |\Psi_k\rangle\,\varepsilon_k\,\langle\Psi_k| \ ; \tag{22.152}$$

Considérant toujours des particules sans interaction, leur Hamiltonien est l'observable additive $H = \sum_i H_1(i)$; l'énergie moyenne E du système dans un état où il y a n_k particules dans l'état $|\Psi_k\rangle$, $k = 1, 2, \ldots$ peut visiblement s'écrire :

$$E = \sum_k n_k \varepsilon_k \ , \tag{22.153}$$

d'où l'expression, en Seconde quantification, du Hamiltonien de ces particules supposées *sans interactions mutuelles* :

$$\boxed{H = \sum_k \varepsilon_k\, N_k \equiv \sum_k \varepsilon_k\, a_k^\dagger a_k} \tag{22.154}$$

De la même façon, l'opérateur donnant le nombre total de particules est :

$$\boxed{N = \sum_k N_k \equiv \sum_k a_k^\dagger a_k} \tag{22.155}$$

Visiblement, N peut s'exprimer à l'aide des opérateurs de champ, ce qui renforce encore l'intérêt de ceux-ci ; en effet, on a :

$$\int_{\mathbb{R}^3} \hat{\psi}^\dagger(\vec{r})\hat{\psi}(\vec{r})\,\mathrm{d}^3 r = \sum_{k,\,k'} \int_{\mathbb{R}^3} \psi_k^*(\vec{r})\psi_{k'}(\vec{r})\, a_k^\dagger a_{k'}\,\mathrm{d}^3 r \ ; \tag{22.156}$$

par l'orthonormalisation des spin-orbitales ψ_k, le second membre est égal à N, d'où à l'envers :

$$\boxed{N = \int_{\mathbb{R}^3} \hat{\psi}^\dagger(\vec{r})\hat{\psi}(\vec{r})\,\mathrm{d}^3 r} \tag{22.157}$$

Bien entendu, il est possible de considérer une autre base d'états à une particule (toujours orthonormalisée), $\{|\Phi_\kappa\rangle\}$, sur laquelle H_1 n'est pas diagonal. Il existe une transformation unitaire permettant d'écrire :

$$|\Psi_k\rangle = \sum_\kappa U_{\kappa k}|\Phi_\kappa\rangle \ , \qquad U_{\kappa k} = \langle\Phi_\kappa|\Psi_k\rangle \ . \tag{22.158}$$

En considérant (22.151), on voit que la même transformation existe entre les opérateurs associés aux deux bases ; avec des notations évidentes :

$$a_k^\dagger = \sum_\kappa U_{\kappa k} A_\kappa^\dagger \ . \qquad (22.159)$$

Cela étant, H (22.154) se transforme comme suit :

$$H = \sum_{\kappa\,\lambda} \sum_k \varepsilon_k U_{\kappa k} U_{\lambda\, k}^* \, A_\kappa^\dagger A_\lambda \ ; \qquad (22.160)$$

la somme sur k est :

$$\sum_k \varepsilon_k \langle \Phi_\kappa | \Psi_k \rangle \langle \Phi_\lambda | \Psi_k \rangle^* = \sum_k \langle \Phi_\kappa | \Psi_k \rangle \, \varepsilon_k \, \langle \Psi_k | \Phi_\lambda \rangle \equiv \langle \Phi_\kappa | H_1 | \Phi_\lambda \rangle \ , \qquad (22.161)$$

de sorte que H s'écrit :

$$H = \sum_{\kappa\,\lambda} \langle \Phi_\kappa | H_1 | \Phi_\lambda \rangle \, A_\kappa^\dagger A_\lambda \ ; \qquad (22.162)$$

D'où la forme générale d'une observable additive à un corps $\Omega = \sum_i \Omega_1(i)$, exprimée en fonction des opérateurs de création et d'annihilation relatifs à des modes quelconques :

$$\boxed{\Omega = \sum_{k\,l} \langle \Psi_k | \Omega_1 | \Psi_l \rangle \, a_k^\dagger a_l \equiv \sum_{k\,l} \omega_{kl} \, a_k^\dagger a_l} \qquad (22.163)$$

Sur ses modes propres, Ω est diagonal et la somme ne contient alors que les termes $k = l$, ce qui fait réapparaître les nombres d'occupation (de ces modes propres), tout comme pour H en (22.154). Revenant à la définition (22.148) des opérateurs de champ, on voit que Ω peut aussi s'exprimer naturellement à l'aide de ceux-ci :

$$\boxed{\Omega = \int_{\mathbb{R}^3} \hat{\psi}^\dagger(\vec{r}) \Omega_1(\vec{r}) \hat{\psi}(\vec{r}) \, \mathrm{d}^3 r} \qquad (22.164)$$

À titre d'illustration, montrons à nouveau en détail comment fonctionne le formalisme. Partons de la quantité $\langle \Psi_k | \Omega | \Psi_l \rangle$, élément de matrice d'une observable additive à un corps entre deux états à une particule engendrés à partir du vide selon (22.105). On peut de ce fait écrire, tenant compte de (22.163) :

$$\langle \Psi_k | \Omega | \Psi_l \rangle = \langle \mathrm{vac} | a_k \, \Omega \, a_l^\dagger | \mathrm{vac} \rangle = \sum_{k'\,l'} \langle \mathrm{vac} | a_k \, a_{k'}^\dagger a_{l'} \, a_l^\dagger | \mathrm{vac} \rangle \, \omega_{k'l'} \ . \qquad (22.165)$$

L'idée consiste à jouer encore avec les relations de commutation[40] afin de mettre les opérateurs d'annihilation à droite dans le produit, juste devant $|\mathrm{vac}\rangle$: en vertu de (22.103), le terme correspondant sera nul. Le produit des quatre opérateurs dans (22.165)

[40] L'opération qui consiste à faire commuter les opérateurs d'un produit pour réduire de deux unités le nombre de facteurs s'appelle une *contraction*.

est égal à $a_k a_{k'}^\dagger(\delta_{ll'} + \varepsilon a_l^\dagger a_{l'})$, avec $\varepsilon = +1$ pour les bosons, -1 pour les fermions ; compte tenu de $a_{l'}|\text{vac}\rangle = 0$, il reste :

$$\langle\Psi_k|\Omega|\Psi_l\rangle = \sum_{k'\,l'} \langle\text{vac}|a_k a_{k'}^\dagger \delta_{ll'}|\text{vac}\rangle\, \omega_{k'l'} \ ; \tag{22.166}$$

en recommençant la même opération avec les opérateurs restants, il vient :

$$\langle\Psi_k|\Omega|\Psi_l\rangle = \sum_{k'\,l'} \langle\text{vac}|(\delta_{kk'} + \varepsilon a_{k'}^\dagger a_k)|\text{vac}\rangle\, \delta_{ll'}\, \omega_{k'l'} \ = \sum_{k'\,l'} \delta_{kk'}\delta_{ll'}\, \omega_{k'l'} = \omega_{kl} \ ; \tag{22.167}$$

Quant au premier membre, c'est très précisément :

$$\langle 0,\, 0,\, \ldots 0,\, 1_k,\, 0,\, \ldots|\sum_i \Omega_1(i)|0,\, 0,\, \ldots 0,\, 1_l,\, 0,\, \ldots\rangle \ ; \tag{22.168}$$

tant que l'indice i porte sur une particule qui n'existe pas (parce que le nombre d'occupation est nul dans le *bra* et dans le *ket*), le résultat est nul ; le seul terme non-nul est bien l'élément de matrice de Ω_1 avec une particule dans $|\Psi_k\rangle$ à gauche et dans $|\Psi_l\rangle$ à droite : c'est bien ω_{kl}, en conformité avec (22.167).

Considérons maintenant une observable à deux corps :

$$V = \frac{1}{2}\sum_{i\neq j} v(i,\, j) \ . \tag{22.169}$$

V est par exemple l'interaction entre particules, $V = \frac{1}{2}\sum_{i\neq j} v(\vec{r}_i - \vec{r}_j)$. En généralisation immédiate de (22.163), on écrit[41] :

$$V = \frac{1}{2}\sum_{k,l,k',l'} v_{kk'll'}\, a_{k'}^\dagger a_k^\dagger a_l a_{l'} \tag{22.170}$$

et toute la question est d'identifier la quantité $v_{kk'll'}$. Pour cela, on fait jouer les contractions comme ci-dessus, en partant de la quantité $\langle\Psi_m\Psi_n|V|\Psi_{m'}\Psi_{n'}\rangle$. Un calcul un peu plus long montre que $v_{kk'll'}$ n'est autre qu'un élément de matrice à deux corps, donné par l'intégrale :

$$v_{kk'll'} = \int \Psi_k^*(\vec{r})\Psi_{k'}^*(\vec{r}\,')v(\vec{r} - \vec{r}\,')\Psi_l(\vec{r})\Psi_{l'}(\vec{r}\,')\mathrm{d}^3r\, \mathrm{d}^3r' \equiv \langle\Psi_k\Psi_{k'}|v|\Psi_l\Psi_{l'}\rangle \tag{22.171}$$

Au total, un Hamiltonien ne contenant que des termes à un et à deux corps :

$$H = \sum_{i=1}^N H_1(i) + \frac{1}{2}\sum_{i=1}^N \sum_{j=1,\, j\neq i}^N H_2(i,\, j) \tag{22.172}$$

s'écrit comme suit :

$$\boxed{H = \sum_{k,l} \varepsilon_{kl}\, a_k^\dagger a_l + \frac{1}{2}\sum_{k,l,k',l'} v_{kk'll'}\, a_{k'}^\dagger a_k^\dagger a_l a_{l'}} \tag{22.173}$$

[41] Attention à l'ordre des indices dans $v_{kk'll'}$ et dans le produit des quatre opérateurs !

où :

$$\varepsilon_{kl} = \langle \Psi_k | H_1 | \Psi_l \rangle \ , \qquad v_{kk'll'} = \langle \Psi_k \Psi_{k'} | H_2 | \Psi_l \Psi_{l'} \rangle \ . \tag{22.174}$$

Remarquer que l'expression (22.173) ne contient pas le nombre (fixe ou variable) de particules, N. Enfin, en terme des opérateurs de champ, le Hamiltonien a le plus souvent typiquement la forme[42] :

$$\boxed{H = \int_{\mathbb{R}^3} \mathrm{d}^3 r \, \hat{\psi}^\dagger(\vec{r}) H_1(\vec{r}, \vec{p}) \hat{\psi}(\vec{r}) + \frac{1}{2} \int_{\mathbb{R}^3} \mathrm{d}^3 r \int_{\mathbb{R}^3} \mathrm{d}^3 r' \, \hat{\psi}^\dagger(\vec{r}) \hat{\psi}^\dagger(\vec{r}\,') H_2(\vec{r}, \vec{r}\,') \hat{\psi}(\vec{r}) \hat{\psi}(\vec{r}\,')} \tag{22.175}$$

22.6 Exercices et problèmes

22.6.1 Retour sur le trou de Fermi

Soit deux fermions de spin 1/2 dont $\vec{S}$ désigne le spin total.

1. A partir de deux orbitales orthonormées, ϕ_1 et ϕ_2, former toutes les fonctions d'onde possibles, propres de $(\vec{S}^2, S_z)$. On notera $\alpha(m_s)$ et $\beta(m_s)$ les deux fonctions définies en (20.36). Que se passe-t-il si $\phi_1 = \phi_2$?
2. Pour tous les états obtenus, déterminer :
   (a) la densité de probabilité $\rho_2(\vec{r}, \vec{r}\,')$ de trouver une particule près de $\vec{r}$ et l'autre près de $\vec{r}\,'$, quels que soient les spins ;
   (b) la densité de probabilité $\rho_1(\vec{r})$ de trouver une particule près de $\vec{r}$.

22.6.2 États de spin de trois électrons

Pour trois électrons :

1. Quelles sont les valeurs possibles du spin total ?
2. Exprimer à l'aide des fonctions α et β toutes les composantes du quadruplet et des doublets.
3. Écrire toutes les fonctions d'onde, propres de $(\vec{S}^2, S_z)$, que l'on peut construire avec trois orbitales ϕ_1, ϕ_2 et ϕ_3.
4. Pour quel(s) état(s) retrouve-t-on une séparation espace - spin après antisymétrisation ?
5. Ces états sont-ils intriqués ?

[42] Pour des particules de spin non-nul, il arrive assez souvent que l'interaction H_2 dépende du spin.

22.6.3 Étude détaillée du groupe des permutations $\mathcal{S}_3$

Soit $\{P_\lambda\}_\lambda$ l'ensemble des permutations de trois objets, muni de la loi de composition interne (22.27) le dotant de la structure de groupe[43] ; P_1 note l'élément neutre (la permutation identité).

1. Dresser la table 6×6 de cette loi de composition.
2. La loi de composition définit une application f de $[1, 6] \times [1, 6]$ dans $[1, 6]$:
$$P_\nu = P_\lambda P_\mu \quad \Longleftrightarrow \quad (\lambda, \mu) \xrightarrow{f} \nu = f(\lambda, \mu) \ ; \tag{22.176}$$
sachant que $\nu = f(\lambda, \mu)$, montrer que $f(\nu, \sigma) = f(\lambda, f(\mu, \sigma))$.
3. On définit une relation d'équivalence $\mathcal{R}$ par :
$$\forall P_\lambda \in \mathcal{S}_3, \forall P_\mu \in \mathcal{S}_3 \ : P_\lambda \mathcal{R} P_\mu \quad \Longleftrightarrow \quad \exists P_\sigma \in \mathcal{S}_3, \ P_\lambda = P_\sigma P_\mu P_\sigma^{-1} \ , \tag{22.177}$$
Former les classes d'équivalence.
4. Soit l'espace vectoriel sur $\mathbb{C}$, $\mathcal{A}(\mathcal{S}_3) \stackrel{\text{déf}}{=} \{X\}$, dont les P_λ constituent une base : $X = \sum_\lambda x_\lambda P_\lambda$, $x_\lambda \in \mathbb{C}$. Grâce à la loi de groupe de $\mathcal{S}_3$, on peut définir une deuxième loi de composition interne dans $\mathcal{A}(\mathcal{S}_3)$:
$$X = \sum_\mu x_\mu P_\mu \ , \ Y = \sum_\nu y_\nu P_\nu \ , Z \stackrel{\text{déf}}{=} XY = \sum_\mu \sum_\nu x_\mu y_\nu P_{f(\mu,\nu)} \ . \tag{22.178}$$
Montrer que $(X+Y)Z = XZ + YZ$, $Z(X+Y) = ZX + ZY$, $(XY)Z = X(YZ)$, $\forall \alpha \in \mathbb{C}$, $X(\alpha Y) = \alpha XY$. $\mathcal{A}(\mathcal{S}_3)$ est alors par définition une *algèbre associative* : c'est l'algèbre du groupe $\mathcal{S}_3$.
5. La deuxième loi interne de l'algèbre étant linéaire, on peut écrire :
$$\forall P_\lambda \in \mathcal{S}_3, \forall X \in \mathcal{A}(\mathcal{S}_3), \ X = \sum_\mu x_\mu P_\mu, \ P_\lambda X = \sum_\mu x_\mu P_{f(\lambda,\mu)} \ . \tag{22.179}$$
À chaque P_λ, on peut donc associer une matrice 6×6 notée $M(P_\lambda)$. Montrer que cet ensemble de matrices a une structure de groupe quand on choisit comme loi de composition interne la multiplication habituelle des matrices.
6. Montrer que ce groupe de matrices constitue une représentation (linéaire) du groupe $\mathcal{S}_3$. La représentation ainsi définie s'appelle la *représentation régulière*, notée Γ^{reg} : c'est celle dont l'espace support est l'algèbre du groupe.

[43] Ce problème fait appel, très intuitivement, à quelques notions de la Théorie de la représentation linéaire des groupes, et constitue une incitation à en savoir plus sur cette question. On pourra consulter à ce sujet l'ouvrage de Bacry [50].

7. Déterminer les valeurs et vecteurs propres des $M(P_\lambda)$. Montrer que ces matrices ont en commun deux vecteurs propres ; en déduire l'existence de deux représentations irréductibles à une dimension $\Gamma^{\rm S}$ et $\Gamma^{\rm A}$; en donner les caractères, $\chi^{\rm I}(P_\lambda)$ (I = A et S).

8. Les deux vecteurs propres trouvées en 7 sont de la forme :

$$Y_{\rm S} = \sum_\lambda \chi^{\rm S}(P_\lambda)P_\lambda \ , \quad Y_{\rm A} = \sum_\lambda \chi^{\rm A}(P_\lambda)P_\lambda \ . \tag{22.180}$$

Calculer $P_\lambda Y_{\rm I}$. En déduire les propriétés de permutation de $Y_{\rm I}\psi(\vec{x}_1, \vec{x}_2, \vec{x}_3)$ où ψ est une fonction quelconque.

9. Les *idempotents primitifs* d'une algèbre, $\{e_i\}_i$, sont définis par les relations :

$$e_i e_j = 0 \ \forall\, i \neq j \ , \quad e_i^2 = e_i \ .$$

Calculer les idempotents primitifs associés aux $Y_{\rm I}$.

10. Soit :

$$X_{11} = P_1 + j^2 P_5 + jP_6 \ , \quad X_{12} = P_2 + jP_3 + j^2 P_4 \ , \tag{22.181}$$

$$X_{22} = P_1 + jP_5 + j^2 P_6 \ , \quad X_{21} = P_2 + j^2 P_3 + jP_4 \ , \tag{22.182}$$

où 1, j, j^2 sont les racines cubiques de l'unité. Montrer que X_{11} et X_{12} d'une part, X_{21} et X_{22} d'autre part, engendrent un espace invariant irréductible. Montrer que les deux représentations correspondantes sont équivalentes. Donner les caractères des P_λ dans ce type de repésentation.

Compte tenu de la forme des matrices des P_λ exprimées sur la base $Y_{\rm S}$, $Y_{\rm A}$, X_{11}, X_{12}, X_{22} et X_{21}, on peut écrire :

$$\Gamma^{\rm reg} = \Gamma^{\rm S} \oplus \Gamma^{\rm A} \oplus 2\Gamma \ . \tag{22.183}$$

Ce résultat est général : dans la décomposition de la représentation régulière d'un groupe fini, chaque représentation irréductible apparaît un nombre de fois égal à sa dimension.

11. Calculer les idemptotents primitifs e_{ii} associés aux X_{ii} ; montrer que

$$P_1 = e_{\rm S} + e_{\rm A} + e_{11} + e_{22} \ .$$

12. On définit le *vecteur-caractère* d'une représentation comme le vecteur dont les composantes sont égales aux caractères des opérations du groupe dans cette représentation. Montrer que les vecteurs-caractères des représentation irréductibles de $\mathcal{S}_3$ sont orthogonaux (ce résultat est général).

13. *Application* : soit un système de trois électrons, et soit σ_{SM} une fonction propre du spin total :

$$\vec{S}^{\,2}\sigma_{SM} = \hbar^2 S(S+1)\sigma_{SM} \ , \qquad S_z\sigma_{SM} = M\hbar\,\sigma_{SM} \ . \tag{22.184}$$

Montrer que $P_\lambda\sigma_{SM}$ est encore fonction propre de $\vec{S}^{\,2}$, S_z, $\forall\, P_\lambda \in \mathcal{S}_3$. Montrer aussi que les fonctions quadruplets se transforment suivant $\Gamma^{\rm S}$ et les fonctions doublets suivant Γ. En déduire que $X_{11}\sigma_{\frac{1}{2}M}$ et que $X_{22}\sigma_{\frac{1}{2}M}$ sont deux états doublets indépendants.

22.6.4 N fermions

Soit $\{\psi_k\}_k$ un ensemble de spin-orbitales orthonormées et Y_A le projecteur antisymétriseur. Un état pour N fermions sans interactions peut être écrit :

$$\Psi = C_\mathrm{A}\, Y_\mathrm{A} \prod_{k=1}^{N} \psi_k \equiv C_\mathrm{A}\, \frac{1}{N!}\, \mathrm{Det}[\psi_1\psi_2 \ldots \psi_N] \ , \tag{22.185}$$

1. Calculer la constante de normalisation C_A.
2. Montrer que deux tels états Ψ sont orthogonaux s'ils diffèrent par au moins une spin-orbitale.
3. Dans la suite, on suppose que N est pair et que les électrons sont appariés deux par deux ; le déterminant de Slater est donc construit avec $N/2$ orbitales $\phi_k(\vec{r}\,)$.

   La densité $D(\vec{r}\,)$ de particules au point $\vec{r}$ dans l'état Ψ est la valeur moyenne de l'opérateur :

   $$\Delta(\vec{r}\,) = \sum_{k=1}^{N} \delta(\vec{r} - \vec{r}_k) \ . \tag{22.186}$$

   Calculer $D(\vec{r}\,) = \langle\Psi|\Delta|\Psi\rangle$. Combien vaut $\int_{\mathbb{R}^3} D(\vec{r}\,)\, \mathrm{d}^3 r$?
4. On note $\vec{x}_k \equiv (\vec{r}_k, \sigma_k)$ l'ensemble des degrés de liberté d'un électron, $\sigma = \pm 1$ désignant la projection de son spin. Soit $\rho(\{\vec{x}_k\}, \{\vec{x}\,'_k\})$ le noyau de l'opérateur densité construit avec le cas pur $|\Psi\rangle$, et soit les deux opérateurs densité *réduits* à une et deux particules :

   $$\gamma(\vec{x}_1,\, \vec{x}\,'_1) \stackrel{\text{déf}}{=} \sum_{\vec{x}_2, \ldots \vec{x}_N} \Psi(\vec{x}_1,\, \vec{x}_2, \ldots, \vec{x}_N)\Psi^*(\vec{x}\,'_1,\, \vec{x}_2, \ldots, \vec{x}_N) \ . \tag{22.187}$$

   $$\Gamma(\vec{x}_1,\, \vec{x}_2,\, \vec{x}\,'_1,\, \vec{x}\,'_2) \stackrel{\text{déf}}{=} \sum_{\vec{x}_3, \ldots \vec{x}_N} \Psi(\vec{x}_1,\, \vec{x}_2,\, \vec{x}_3 \ldots, \vec{x}_N)\Psi^*(\vec{x}\,'_1,\, \vec{x}\,'_2,\, \vec{x}_3,\, \ldots, \vec{x}_N) \ . \tag{22.188}$$

   (a) Quelles sont les traces de γ et Γ ?

   (b) Soit $\rho_\mathrm{e}(\vec{r},\, \vec{r}\,') \stackrel{\text{déf}}{=} N^{-1} \sum_\sigma \gamma(\vec{r},\, \sigma,\, \vec{r}\,', \sigma)$. Si l'on interprète ρ_e comme la matrice d'un électron dans le *bain* des $N-1$ autres, s'agit-il d'un cas pur ou d'un cas mixte ?

   (c) Quelle est l'entropie de Shannon - Wiener[44] de ρ_e ?
5. Le Hamiltonien des fermions sans interaction est de la forme :

   $$H(1,\, 2,\, \ldots N) = \sum_{k=1}^{N} H_1(k) \ ; \tag{22.189}$$

   les états propres de H_1 sont supposés connus, $H_1\psi_n = E_n\psi_n$; calculer $\langle\Psi|H|\Psi\rangle$.

[44] Voir note 10, p. 923.

6. Soit l'interaction à N-corps :

$$V(1, 2, \dots N) = \sum_{k=1}^{N} \sum_{k>l}^{N} v(k, l)) \tag{22.190}$$

où v est un terme d'interaction à deux particules. Montrer que $\langle\Psi|V|\Psi\rangle$ s'exprime comme une somme de termes directs et d'échange.

7. Expliquer comment toutes les valeurs moyennes d'observables du type H et V peuvent être calculées à partir des *seuls* opérateurs réduits définis en 4.

22.6.5 N fermions libres

On reprend le problème 22.6.4 en supposant que les fermions ne sont soumis à aucun champ de force et enfermés dans une boîte cubique de côté L. On prend des conditions aux limites périodiques, de sorte que les orbitales sont des ondes planes :

$$\phi_{\vec{k}}(\vec{r}) = L^{-3/2}\, \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \;; \tag{22.191}$$

on fera à chaque fois en temps utile la substitution $\sum_{\vec{k}} \rightarrow (\frac{L}{2\pi})^3 \int \mathrm{d}^3 k$.

1. Soit k_{F} le module du vecteur d'onde $\vec{k}$ associé à la plus haute orbitale occupée ; exprimer k_{F} en fonction de la densité fermionique.

2. La densité réduite est :

$$\gamma(\vec{r}, \vec{r}\,') = 2 \sum_{k<k_{\mathrm{F}}} \phi_{\vec{k}}(\vec{r})\phi^*_{\vec{k}}(\vec{r}\,') \;; \tag{22.192}$$

montrer qu'elle s'exprime à l'aide d'une fonction de Bessel sphérique (voir (19.83)).

3. Que représente $\gamma(\vec{r}, \vec{r})$?

4. On définit comme suit la fonction de corrélation de paire, $C(\vec{r}, \vec{r}\,')$:

$$C(\vec{r}, \vec{r}\,') \stackrel{\text{déf}}{=} \frac{1}{2}\gamma(\vec{r}, \vec{r})\gamma(\vec{r}\,', \vec{r}\,') - \frac{1}{4}[\gamma(\vec{r}, \vec{r}\,')]^2 \;; \tag{22.193}$$

écrire son expression et tracer l'allure de son graphe en fonction de $k_{\mathrm{F}}r$, où r désigne la distance $\|\vec{r} - \vec{r}\,'\|$.

22.6.6 Correction quantique à la fonction de partition classique d'un gaz parfait

Soit dans $\mathbb{R}^3$ un gaz parfait de $N \gg 1$ particules identiques (fermions ou bosons), de masse m et de spin S. Un état du gaz se forme à partir d'un produit $\Phi \stackrel{\text{déf}}{=} \prod_{i=1}^{N} \psi_{a_i}$ de spin-orbitales orthonormalisées, où a_i note collectivement l'ensemble des nombres quantiques

nécessaires pour spécifier complètement l'état d'une particule. Dans ces conditions, tout état satisfaisant le postulat d'antisymétrisation est $|\Psi\rangle = N!^{1/2}\, Y|\Phi\rangle$, où Y est le bon projecteur $Y_{\rm S}$ ou $Y_{\rm A}$ suivant qu'il s'agit de bosons ou de fermions.

1. Montrer que $\langle\Psi|\Psi\rangle = \prod_{a_i} n_{a_i}!$

2. L'opération de trace (Tr) consiste à sommer sur tous les états distincts du gaz ; montrer que pour une observable quelconque Ω, on a :

$$\mathrm{Tr}\,\Omega = \sum_{a_1}\sum_{a_2}\ldots\sum_{a_N}\langle\Phi|\Omega Y|\Phi\rangle \ , \qquad (22.194)$$

   où chaque sommation se fait *indépendamment* des autres.

3. La fonction de partition canonique est par définition $Z \stackrel{\text{déf}}{=} \mathrm{Tr}\,\mathrm{e}^{-\beta H}$ ($\beta = 1/k_{\rm B}T$). Quelle est ici la forme de H ?

4. Les orbitales étant des ondes planes, on les normalise dans une boîte cubique de côté L, aussi grand que l'on veut : $\psi_{\vec{k}}(\vec{r}) = L^{-3/2}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$; les spin-orbitales sont notées $\psi_{\vec{k},M}$, avec $M = -S, -S+1, ..., +S$, et sont représentables par $\psi_{\vec{k}}(\vec{r}) \otimes |M\rangle$ où $|M\rangle$ est une matrice-colonne $1 \times \mathrm{g}$ ayant des zéros partout sauf dans la $M^{\rm e}$ ligne où l'élément de matrice vaut 1 ($\mathrm{g} \stackrel{\text{déf}}{=} 2S+1$).

   En effectuant la sommation sur les vecteurs d'onde ($\sum_{\vec{k}} \to \left(\frac{L}{2\pi}\right)^3 \int \mathrm{d}^3 k$), montrer que la fonction de partition $Z(T,\, V,\, N)$ est donnée par l'expression ($V = L^3$) :

$$Z(T,\, V N) = \frac{1}{N!}\sum_{\text{spins}}\left(\frac{2\pi m}{\beta h^2}\right)^{3N/2}\int_{\mathbb{R}^{3N}}\prod_{n=1}^{N}\mathrm{d}^3 r_n\Big[1 \pm \sum_{i<j}\delta_{M_i M_j}\mathrm{e}^{-\frac{m}{\beta\hbar^2}(\vec{r}_i-\vec{r}_j)^2} + \ldots\Big] \ , \qquad (22.195)$$

   où les signes + et − se réfèrent respectivement aux bosons et aux fermions, et où les ... désignent des termes venant de permutations de plus de deux objets ; ces termes sont omis dans toute la suite.

5. En déduire que la fonction de partition (approchée) a pour expression :

$$Z(T,\, V N) = \frac{1}{N!}\Big(\mathrm{g}\frac{V}{\lambda_T^3}\Big)^N\Big[1 \pm \frac{N}{2^{5/2}\mathrm{g}}\Big(\frac{\lambda_T}{d}\Big)^3\Big] \ , \qquad (22.196)$$

   où $d \stackrel{\text{déf}}{=} (V/N)^{1/3}$ est la distance moyenne entre particules et $\lambda_T \stackrel{\text{déf}}{=} \frac{h}{\sqrt{2\pi m k_{\rm B}T}}$ la longueur d'onde thermique de de Broglie.

6. Montrer[45] que la limite thermodynamique ($N \to +\infty$, $V \to +\infty$, $\frac{N}{V} = \mathrm{C^{ste}} \stackrel{\text{déf}}{=} \rho$) de $\big[Z(T,\, V N)\big]^{1/N}$ est :

$$Z_\infty(T,\, \rho) \stackrel{\text{déf}}{=} \lim \text{thermo}\big[Z(T,\, V N)\big]^{1/N} = \mathrm{eg}\Big(\frac{d}{\lambda_T}\Big)^3\Big[1 \pm \frac{1}{2^{5/2}\mathrm{g}}\Big(\frac{\lambda_T}{d}\Big)^3\Big] \ , \qquad (22.197)$$

   Comparer ce résultat avec l'expression classique, obtenue directement.

[45] Utiliser la formule de Stirling (voir p. 816).

7. Établir la condition sur la densité ρ exprimant la séparation entre les régimes classique et quantique. Faire les applications numériques pour un gaz d'hydrogène (atomique) et un gaz d'électrons. Qu'en conclure physiquement suivant qu'il s'agit des électrons d'un métal ou d'une vapeur atomique dans des conditions normales ?

8. Montrer que l'on peut interpréter la première correction quantique comme émanant d'un potentiel effectif à deux corps, $W(\|\vec{r}_1 - \vec{r}_2\|)$ dépendant de la température, attractif pour les bosons, répulsif pour les fermions. Tracer $W(r)$ en fonction de r.

9. Quelle est la correction quantique à l'énergie libre de Helmoltz par particule, F ?

10. En déduire que, dans cette approximation et en dépit de toute interaction au sens classique du terme, l'énergie du gaz satisfait :

$$E_{\text{bosons}} < E_{\hbar=0} < E_{\text{fermions}} \quad . \tag{22.198}$$

11. Examiner la première correction quantique pour quelques grandeurs thermodynamiques.

22.6.7 Équations du mouvement pour les opérateurs de champ

Le Hamiltonien d'un ensemble de particules identiques (bosons ou fermions) est écrit sous la forme :

$$H = \sum_{k,l} a_k^\dagger \langle k|H_0|l\rangle a_l + \frac{1}{2} \sum_{k,l,k',l'} a_{k'}^\dagger a_k^\dagger \langle k\,k'|V|l\,l'\rangle a_l a_{l'} \tag{22.199}$$

1. Établir les relations :

$$[a_j,\, a_k^\dagger a_l] = a_l \delta_{kj} \quad , \qquad [a_j,\, a_q^\dagger a_r^\dagger a_s a_t] = a_r^\dagger a_s a_t \delta_{qj} + a_q^\dagger a_t a_s \delta_{rj} \quad . \tag{22.200}$$

2. Écrire l'équation de Heisenberg pour l'opérateur $a_j(t)$.

3. Utiliser la propriété de symétrie $\langle k\,k'|V|l\,l'\rangle = \langle k'\,k|V|l'\,l\rangle$ pour montrer que cette équation peut être mise sous la forme :

$$\mathrm{i}\hbar \frac{\mathrm{d}a_j(t)}{\mathrm{d}t} = \sum_l \langle j|H_0|l\rangle\, a_j(t) + \sum_{k,\,l,\,m} a_k^\dagger(t) a_l(t) a_m(t) \langle jk|V|ml\rangle \quad . \tag{22.201}$$

4. On choisit la représentation où l'interaction est diagonale suivant :

$$\langle jk|V|ml\rangle = V_{jk} \delta_{jm} \delta_{kl} \quad ; \tag{22.202}$$

on note $\hat{\psi}_\sigma(\vec{r},\, t)$ les opérateurs de champ, σ désignant la projection du spin. Dans l'hypothèse où l'interaction ne dépend pas du spin, montrer que ces opérateurs ont pour équation du mouvement :

$$\mathrm{i}\hbar\frac{\partial\hat{\psi}_\sigma(\vec{r},\, t)}{\partial t} = \sum_{\sigma'}\int\langle\vec{r}\sigma\,|H_0|\vec{r}\,'\sigma'\rangle\hat{\psi}_{\sigma'}(\vec{r}\,',\, t)\,\mathrm{d}^3r' + \sum_{\sigma'}\int V(\vec{r},\, \vec{r}\,')\hat{\psi}_{\sigma'}(\vec{r}\,',\, t)^\dagger\hat{\psi}_{\sigma'}(\vec{r}\,',\, t)\hat{\psi}_\sigma(\vec{r},\, t)\,\mathrm{d}^3r' \ . \quad (22.203)$$

5. Cette équation est-elle linéaire ? Interpréter le terme d'interaction en tant que potentiel effectif, et suggérer une méthode itérative de résolution.

6. Comme c'est souvent le cas, le Hamiltonien à une particule H_0 est supposé être une fonction des seules position et impulsion ; en représentation-q, on écrit alors :

$$\langle\vec{r}\sigma\,|H_0|\vec{r}\,'\sigma'\rangle = H_0(\vec{r},\, -\mathrm{i}\hbar\vec{\nabla})\,\delta(\vec{r}-\vec{r}\,')\,\delta_{\sigma\sigma'} \ ; \quad (22.204)$$

en déduire l'équation du mouvement pour $\hat{\psi}_\sigma(\vec{r},\, t)$, vraie qu'il s'agisse de bosons ou de fermions.

7. Que devient l'équation précédente en l'absence d'interaction entre les particules ? Commenter.

physique

Chapitre 23

Méthodes d'approximation pour les états propres

Le but de ce chapitre est de présenter deux méthodes d'approximation d'usage courant pour les états propres liés d'un Hamiltonien que l'on ne sait pas traiter exactement : la méthode variationnelle et la théorie des perturbations stationnaires.

23.1 Méthode variationnelle

23.1.1 Formulation variationnelle de l'équation aux valeurs et vecteurs propres

L'équation aux valeurs et fonctions propres :

$$H\,\psi(\vec{r}) = E\,\psi(\vec{r}) \ , \tag{23.1}$$

apparaît après la séparation espace - temps dans l'équation de Schrödinger, ayant posé $\Psi(\vec{r}, t) = \mathrm{e}^{\frac{1}{i\hbar}Et}\psi(\vec{r})$, et peut être considérée comme l'équation d'Euler - Lagrange d'un problème variationnel[1]. En effet, soit à trouver la fonction $\psi(\vec{r})$ satisfaisant la condition de normalisation :

$$\int_{\mathbb{R}^3} |\psi(\vec{r})|^2 \,\mathrm{d}^3 r = 1 \ , \tag{23.2}$$

[1]Historiquement, c'est d'ailleurs sous la forme variationnelle que Schrödinger a d'abord introduit son équation [67] dans le premier [175] de ses quatre articles fondateurs de la Mécanique quantique.

Une belle présentation de la méthode d'Euler - Lagrange peut être trouvée dans l'ouvrage classique de Goldstein [51].

et rendant stationnaire l'expression suivante :

$$\int_{\mathbb{R}^3} \Big[-\frac{\hbar^2}{2m}\psi^*(\vec{r})\Delta\psi(\vec{r}) + \psi^*(\vec{r})V(\vec{r})\psi(\vec{r})\Big]\,\mathrm{d}^3r \ , \tag{23.3}$$

qui n'est autre que la valeur moyenne du Hamiltonien H :

$$H = \frac{\vec{p}^{\,2}}{2m} + V(\vec{r}) \tag{23.4}$$

dans l'état normalisé ψ, que l'on peut noter $\langle H\rangle$ ou, plus conformément à l'usage dans un contexte variationnel, $E[\psi]$; les crochets rappellent qu'il s'agit d'une *fonctionnelle* de la fonction ψ : E est un *nombre* (réel) attaché à une certaine famille de *fonctions*.

Avec les conditions habituelles (annulation de ψ aux bornes de l'intervalle, ce qui est toujours réalisé pour les états liés), une intégration par parties transforme l'intégrale (23.3) en[2] :

$$E[\psi] = \int_{\mathbb{R}^3} \Big[+\frac{\hbar^2}{2m}\,\vec{\nabla}\psi^*(\vec{r}).\vec{\nabla}\psi(\vec{r}) + \psi^*(\vec{r})V(\vec{r})\psi(\vec{r})\Big]\,\mathrm{d}^3r \ . \tag{23.6}$$

On va montrer que la fonctionnelle $E[\psi]$ est extrémale pour les états propres du Hamiltonien ; pour la simplicité de l'écriture, la démonstration est faite à une dimension d'espace.

Soit donc à rendre extrémale la fonctionnelle :

$$E[\psi] = \int_{-\infty}^{+\infty} \Big[+\frac{\hbar^2}{2m}\,\psi'^*(x)\,\psi'(x) + \psi^*(x)V(x)\psi(x)\Big]\,\mathrm{d}x \tag{23.7}$$

avec la condition :

$$\int_{-\infty}^{+\infty} \psi^*(x)\psi(x)\,\mathrm{d}x = 1 \ . \tag{23.8}$$

La variation d'une fonctionnelle avec contrainte(s) se fait commodément en introduisant un (des) multiplicateur(s) de Lagrange [51] – qui est (sont) autant d'inconnue(s) supplémentaire(s) à déterminer – mais qui permettent un calcul plus souple en général et surtout plus symétrique. Il y a ici une seule contrainte – exprimée par (23.8) – donc un seul multiplicateur, noté λ. La fonctionnelle à rendre extrémale est alors :

$$F[\psi] \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} \Big[+\frac{\hbar^2}{2m}\,\psi'^*(x)\,\psi'(x) + \psi^*(x)V(x)\psi(x) \ - \ \lambda\psi^*(x)\psi(x)\Big]\,\mathrm{d}x$$
$$\equiv \int_{-\infty}^{+\infty} \Phi(\psi, \psi', \psi^*, \psi'^*)\,\mathrm{d}x \ . \tag{23.9}$$

[2] L'expression (23.6) montre que la valeur moyenne de l'énergie cinétique est :

$$E_{\text{cin}}[\psi] = \frac{\hbar^2}{2m}\int_{\mathbb{R}^3} |\vec{\nabla}\psi(\vec{r})|^2\,\mathrm{d}^3r \ , \tag{23.5}$$

et il saute aux yeux que cette quantité est positive, comme il se doit.

Oublions provisoirement que les fonctions ψ et ψ^* sont complexes conjuguées l'une de l'autre – ce qui semble rendre leurs variations non-indépendantes. En considérant ψ et ψ^* comme indépendantes, on peut écrire deux équations d'Euler-Lagrange[3] pour la fonctionnelle $F[\psi] \stackrel{\text{déf}}{=} \int \Phi \, \mathrm{d}x$:

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial \Phi}{\partial \psi'^*} - \frac{\partial \Phi}{\partial \psi^*} = 0 \ , \qquad \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial \Phi}{\partial \psi'} - \frac{\partial \Phi}{\partial \psi} = 0 \ , \tag{23.10}$$

qui s'explicitent comme suit :

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\hbar^2}{2m}\psi' - V\psi + \lambda\psi = 0 \ , \qquad \frac{\mathrm{d}}{\mathrm{d}x}\frac{\hbar^2}{2m}\psi'^* - V\psi^* + \lambda\psi^* = 0 \ , \tag{23.11}$$

soit :

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V\psi = \lambda\psi \ , \qquad -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi^*}{\mathrm{d}x^2} + V\psi^* = \lambda\psi^* \ . \tag{23.12}$$

Ces deux équations montrent que le multiplicateur λ n'est autre que l'énergie, qui est réelle. De surcroît, elles se déduisent l'une de l'autre par conjugaison puisque V est réel[4] ; tout est donc cohérent et il apparaît après coup qu'il était parfaitement licite de considérer les variations de ψ et de ψ^* comme indépendantes[5]. Finalement, la condition d'extremum pour $F[\psi]$ est :

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + V\psi = \lambda\psi \ , \tag{23.13}$$

qui n'est autre que l'équation aux valeurs propres (23.1).

♦ *Remarque*

Le calcul a été fait dans le cas où H est de la forme (23.4). Tout ceci se généralise aux cas plus complexes, par exemple lorsqu'un champ magnétique est présent. En outre, la généralisation au cas tridimensionnel est un pur jeu d'écriture. ♦

L'équivalence :

$$\boxed{\delta E[\psi] = 0 \quad \Longleftrightarrow \quad H\psi = E\,\psi} \tag{23.14}$$

est un résultat majeur, qui peut aussi s'obtenir de façon plus formelle ([15], p. 654), en utilisant la notation de Dirac. Soit la fonctionnelle définie sur l'espace des états :

$$E[|\psi\rangle] = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \ . \tag{23.15}$$

---

[3] Les équations (23.10) sont structurellement identiques aux équations de Lagrange, où ψ est remplacé par une coordonnée q et x est le temps ; la fonctionnelle minimisée est alors l'action S et Φ est le Lagrangien L.

[4] V réel assure que l'équation de conservation de la probabilité est satisfaite, que H est hermitique, que l'évolution est unitaire, etc.

[5] Ceci tient au fait que la variation d'une fonction à valeurs complexes implique la variation de deux fonctions (la partie réelle et la partie imaginaire) ; l'annulation de la variation de la fonctionnelle conduit donc de fait à deux équations réelles.

On va montrer que tout $|\psi\rangle$ rendant E stationnaire est vecteur propre du spectre discret de H – et réciproquement –, étant entendu que tout état discret est normalisable, assurant que le dénominateur dans (23.15) existe et est fini ; la valeur propre associée n'est autre que la valeur de la fonctionnelle $E[|\psi\rangle]$ pour l'état propre en question. Afin que la définition de E (23.15) ait un sens, seuls des vecteurs de norme *finie* sont considérés dans la suite.

Quand $|\psi\rangle$ varie de $|\delta\psi\rangle$, $E[|\psi\rangle]$ varie de δE :

$$\delta E = E[|\psi\rangle + |\delta\psi\rangle] - E[|\psi\rangle] \ ; \tag{23.16}$$

L'expression de δE s'obtient comme quand on calcule explicitement la variation d'une fonction ordinaire $f(x)$ quand son argument varie de dx. En calquant sur l'expression de la dérivée d'une fraction, on voit que la variation de $E[|\psi\rangle]$ s'écrit :

$$\delta E[|\psi\rangle] = \frac{\langle\delta\psi|H|\psi\rangle + \langle\psi|H|\delta\psi\rangle}{\langle\psi|\psi\rangle} - \langle\psi|H|\psi\rangle \frac{\langle\delta\psi|\psi\rangle + \langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle^2} + \mathcal{O}(\delta\psi^2) \ , \tag{23.17}$$

soit :

$$\begin{aligned}\langle\psi|\psi\rangle\, \delta E[|\psi\rangle] &= \langle\delta\psi|H|\psi\rangle + \langle\psi|H|\delta\psi\rangle - E[|\psi\rangle]\,(\langle\delta\psi|\psi\rangle + \langle\psi|\delta\psi\rangle) + \mathcal{O}(\delta\psi^2) \\ &= \langle\delta\psi|(H-E)|\psi\rangle + \langle\psi|(H-E)|\delta\psi\rangle + \mathcal{O}(\delta\psi^2) \ . \end{aligned} \tag{23.18}$$

Tant que $\langle\psi|\psi\rangle$ est non nul et fini, la condition d'extremum, $\delta E = 0$, s'écrit en annulant le facteur de l'incrément linéaire en $\delta\psi$:

$$\langle\delta\psi|(H-E)|\psi\rangle + \langle\psi|(H-E)|\delta\psi\rangle = 0 \ . \tag{23.19}$$

À nouveau se pose la question de l'indépendance des deux variations $\langle\delta\psi|$ et $|\delta\psi\rangle$; l'argument antérieur vaut toujours : puisque la théorie s'appuie sur le corps des complexes, il est licite de les considérer comme indépendantes[6]. De fait, remplaçons $|\delta\psi\rangle$ par $\mathrm{i}\,|\delta\psi\rangle$; par l'antilinéarité du produit scalaire, la condition (23.19) devient :

$$-\mathrm{i}\,\langle\delta\psi|(H-E)|\psi\rangle + \mathrm{i}\,\langle\psi|(H-E)|\delta\psi\rangle = 0 \ . \tag{23.21}$$

La combinaison des deux dernières équations montre bien que la condition d'extremum, quelles que soient les deux variations $|\delta\psi\rangle$ et $\langle\delta\psi|$, s'exprime par les deux équations :

$$(H-E)|\psi\rangle = 0 \ , \qquad \langle\psi|(H-E) = 0 \tag{23.22}$$

mais comme H est hermitique, l'une d'entre elles suffit.

Réciproquement, soit $|\psi\rangle$ un état propre de H associé à l'énergie E. Par définition on a $H|\psi\rangle = E|\psi\rangle$, d'où $(H-E)|\psi\rangle = 0$.

[6]Pour se convaincre autrement ([25], p. 1138) de l'indépendance de fait des deux variations, on peut aussi poser $|\delta\psi\rangle = \epsilon|\phi\rangle$ où $|\phi\rangle$ est donné et où ϵ est un infiniment petit réel ; alors $\langle\delta\psi| = \epsilon\langle\phi|$ et la condition devient :

$$2\,\langle\delta\phi|(H-E)|\psi\rangle = 0 \ . \tag{23.20}$$

En résumé, tant qu'il s'agit d'états normalisables, il y a équivalence entre la propriété d'être un état propre et celle de rendre $E[|\psi\rangle]$ extrémale. Cette affirmation est parfois appelée *théorème de Ritz* [25].

Ce théorème est très utilement complété par le corollaire suivant ; pour un système donné, quel que soit l'état $|\phi\rangle$, on a :

$$E[\phi] \geq E_{\text{fond}} \equiv E_1 \ . \tag{23.23}$$

Ceci signifie que, pour tout Hamiltonien, sa valeur moyenne avec n'importe quelle fonction d'essai $|\phi\rangle$ normalisable est toujours plus grande (ou égale) que l'énergie de l'état fondamental exact. Pour établir ce résultat, calculons la différence $E[\phi] - E_{\text{fond}}$ sur la base propre du Hamiltonien du système considéré[7] :

$$\begin{aligned} E[\phi] - E_{\text{fond}} &= \frac{1}{\langle\phi|\phi\rangle}[\langle\phi|H|\phi\rangle - E_1\langle\phi|\phi\rangle] = \frac{1}{\langle\phi|\phi\rangle}\langle\phi|\Big[\sum_n(E_n|E_n\rangle\langle E_n|) - E_1\Big]|\phi\rangle \\ &= \frac{1}{\langle\phi|\phi\rangle}\langle\phi|\Big[\sum_n(E_n - E_1)|E_n\rangle\langle E_n|\Big]|\phi\rangle \end{aligned} \tag{23.24}$$

Le préfacteur est positif (c'est l'inverse du carré d'une norme), ainsi que chaque terme de la série[8], d'où l'inégalité annoncée. On peut aussi raisonner comme suit ; soit le développement :

$$|\phi\rangle = \sum_n c_n|E_n\rangle \ ; \tag{23.25}$$

alors, comme $E_n \geq E_1 \ \forall n$:

$$E[\phi] = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \geq \frac{\sum_n |c_n|^2 E_1}{\sum_n |c_n|^2} \equiv E_1 \ . \tag{23.26}$$

L'inégalité devient une égalité dans le seul cas où tous les c_n sont nuls, sauf c_1 qui vaut alors 1, auquel cas $|\phi\rangle$ n'est autre que l'état *exact* $|E_1\rangle$. Ainsi, quelle que soit la fonction d'essai considérée, la valeur moyenne de l'énergie est toujours supérieure (ou égale) à la valeur fondmentale :

$$\boxed{E[\phi] \geq E_1} \tag{23.27}$$

23.1.2 Calcul variationnel des états discrets (états liés)

L'usage de la méthode variationnelle et la résolution directe de l'équation aux valeurs propres sont donc deux démarches totalement équivalentes. Il en résulte que leur achèvement exact est d'une égale difficulté et, d'une façon générale, quand on ne sait pas trouver les solutions exactes d'une façon, on ne sait pas plus les obtenir par l'autre.

[7] La démonstration suppose pour simplifier que tous les états discrets forment une base complète.

[8] L'opérateur entre crochets dans (23.24) est un opérateur défini positif puisque $E_n \geq E_1$ quel que soit n ; de ce fait, sa valeur moyenne dans n'importe quel état est un nombre positif.

En pratique, la méthode variationnelle est cependant utile pour obtenir des valeurs propres et des fonctions propres approchées. En effet, en vertu du corollaire (23.27), on peut adopter la méthodologie suivante. Une famille de fonctions étant donnée (ce n'est pas forcément un espace vectoriel – elles peuvent dépendre non-linéairement d'un ou plusieurs paramètres), on peut espérer obtenir une approximation convenable de l'état fondamental en choisissant dans cette famille la fonction $|\psi_{\text{opt}}\rangle$ qui donne sa plus *petite* valeur à la fonctionnelle $E[|\psi\rangle]$. Par définition, de toutes les fonctions de cette famille, $|\psi_{\text{opt}}\rangle$ sera celle qui donne l'énergie la plus proche de E_{fond}. Ce critère vaut ce qu'il vaut – voir plus loin.

À titre d'illustration, prenons un exemple, académique certes puisqu'il traite un cas où l'on connaît la solution exacte, mais qui a ses vertus puisqu'il permet précisément de faire une comparaison systématique entre les fonctions d'onde exacte et approchée. Soit à trouver par ce type de variation la fonction d'onde radiale fondamentale de l'atome d'hydrogène. On pose :

$$R(r) = a_0^{-3/2} f(\rho) \qquad (\rho = r/a_0) \tag{23.28}$$

et il s'agit de trouver la meilleure fonction $f(\rho)$ dans une famille à préciser par la suite. S'agissant d'un état s, le terme centrifuge est nul, l'énergie cinétique se réduisant à $\frac{1}{2m}p_r^2$ (voir (19.9)) ; compte tenu de l'expression de p_r^2 (voir (19.22)), la fonctionnelle $E[R]$ est :

$$E[R] = \frac{1}{\int_0^{+\infty} r^2\, R^* R\, \mathrm{d}r} \int_0^{+\infty} r^2\, R^* \left[\frac{-\hbar^2}{2m}\left(R'' + \frac{2}{r}R'\right) - \frac{e'^2}{r}\right] \mathrm{d}r \ , \tag{23.29}$$

et s'écrit à l'aide de f comme suit :

$$E[f] = E_1 \, \frac{\int_0^{+\infty} f^* [\rho^2 f'' + 2\rho f' + 2\rho f]\, \mathrm{d}\rho}{\int_0^{+\infty} \rho^2 f^* f\, \mathrm{d}\rho} \tag{23.30}$$

où E_1 est l'énergie du fondamental de l'hydrogène :

$$E_1 = -\frac{me'^4}{2\hbar^2} = -\frac{e'^2}{2a_0} \simeq -13{,}6\,\text{eV} \ , \qquad a_0 = \frac{\hbar^2}{me'^2} \tag{23.31}$$

Choisissons maintenant des familles de fonctions, étant entendu que le choix est illimité ; pour l'exemple, considérons[9] :

$$f_a(\rho) = \mathrm{e}^{-a\rho} \ , \qquad f_b(\rho) = \frac{1}{\rho^2 + b^2} \tag{23.32}$$

où a et b ont le sens de longueurs adimensionnées (longueurs en unités a_0), constituant pour chaque fonction le paramètre variationnel. Quand celui-ci varie, la fonction correspondante décrit une famille de fonctions qui peut toujours être dotée d'une structure d'espace vectoriel, mais dont la famille ne saurait visiblement constituer ici une base.

[9]Ces fonctions ne sont pas normalisées.

$E[f_a]$ et $E[f_b]$ sont donc ici de simples fonctions d'un paramètre variationnel, a ou b. Le calcul est facile et on trouve :

$$E[f_a] = (2a - a^2)\, E_1 \equiv \varepsilon_a(a) \ , \qquad E[f_b] = \frac{8b - \pi}{2\pi b^2}\, E_1 \equiv \varepsilon_b(b) \ . \tag{23.33}$$

La minimisation donne :

$$a_{\text{min}} = 1, \quad \varepsilon_a(a_{\text{min}}) = E_1 \ , \qquad b_{\text{min}} = \frac{\pi}{4}, \quad \varepsilon_b(b_{\text{min}}) = \frac{8}{\pi^2}\, E_1 \simeq 0,81\, E_1 \ . \tag{23.34}$$

Comme $E_1 < 0$, on a bien $\varepsilon_b(b_{\text{min}}) > E_1$.

En ce qui concerne f_a, on voit que la fonction optimisée $f_{a,\,\text{opt}}$ donne la valeur exacte de l'énergie, c'est-à-dire que cette fonction *est* la fonction exacte (voir le commentaire suivant (23.26)). De fait, la famille choisie contenait la fonction exacte $\mathrm{e}^{-\rho}$ et la méthode des variations est allée la *pêcher* ; le but de l'exercice était de montrer ce phénomène.

En revanche, si elle fournit une valeur raisonnable de l'énergie (avec évidemment $E_{\text{opt}} > E_1$ puisque $E_1 < 0$), la fonction $f_{b,\,\text{opt}}$ est loin d'être satisfaisante et représente très mal la vraie fonction : notamment, elle décroît (à grande distance) très lentement, comme r^{-2}, alors que la décroissance exacte est exponentielle. Ceci donne des résultats fort médiocres : la valeur moyenne exacte de r, distance au noyau, vaut $1,5a_0$ (voir chapitre 19, sous section 19.2.2), alors que toutes les fonctions f_b, (l'optimisée comme les autres), donnent ... $\langle r \rangle = +\infty$ (à l'infini, l'intégrand donnant $\langle r \rangle$ calculé avec f_b varie comme $r^2 r^{-4} r = r^{-1}$: la divergence de $\langle r \rangle$ est logarithmique).

Il est clair sur ce dernier exemple que la valeur de l'énergie ne saurait définir un critère sûr de qualité pour la fonction d'onde ainsi trouvée. Ceci n'est pas surprenant : la valeur de l'énergie variationnelle résulte d'une intégration où des erreurs grossières de comportement de la fonction d'onde peuvent avoir le mauvais goût de se compenser, même si ce n'est pas toujours le cas, donnant finalement pour l'énergie une valeur acceptable en dépit de la médiocrité locale de la fonction.

Une façon plus sûre d'apprécier la qualité d'une fonction variationnelle repose sur une bonne mesure de la proximité des fonctions exacte et approchée ; on peut par exemple penser au produit scalaire S :

$$S \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} \psi_{\text{opt}}^* \, \psi_1 \, \mathrm{d}^3 r \ , \tag{23.35}$$

qui vaut 1 si $\psi_{\text{opt}}^* = \psi_1$, mais qui peut néanmoins avoisiner cette valeur avec une mauvaise fonction ψ_{opt}^* à la suite de compensations malheureuses. Un bien meilleur index est la distance d_n :

$$d_n \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} \|\psi_{\text{opt}} - \psi_1\|^n \, \mathrm{d}^3 r \ , \tag{23.36}$$

avec $n = 1$ ou 2, par exemple : plus d_n est proche de zéro, meilleure est la fonction ψ_{opt}. Évidemment, un tel calcul n'est possible ... que si l'on connaît la fonction exacte ψ_1 !

Le calcul de S ou d_n dans quelques cas d'école permet de réaliser qu'il n'y a pas de corrélation directe entre l'erreur sur l'énergie et l'erreur sur la fonction d'onde telle que S ou d_n permettent de la définir.

Malgré la difficulté pour évaluer la qualité des résultats qu'elle produit, la méthode variationnelle reste une méthode puissante et relativement efficace ; en toute circonstance, on doit se laisser guider, lors du choix de la famille d'essai, par des considérations fondées sur des arguments généraux ou sur des propriétés dont on sait que la fonction exacte inconnue doit les posséder : nombre de nœuds, symétrie, comportement à l'origine et asymptotique, etc. ; d'une façon générale, des considérations physiques sont toujours fort utiles.

Compte tenu de l'inégalité (23.27), la méthode variationnelle convient surtout à la description de l'état fondamental ; en effet, on ne dispose *a priori* d'aucune relation d'ordre analogue pour les états excités et la minimisation peut faire "plonger" l'énergie, loin de l'énergie d'un état excité donné, que l'on veut décrire. On peut même dire que, pour une variation linéaire (i.e. sur les composantes d'un vecteur dans un espace effectif), meilleur est l'état fondamental, plus médiocres sont en général les résultats concernant les états excités, en vertu de l'invariance des traces.

D'un autre côté, dans le cas exceptionnel où l'on connaît le fondamental *exact* $|E_1\rangle$, on peut tout redémontrer en se plaçant d'emblée dans l'espace orthogonal à ce vecteur ; en particulier :

$$E[|\phi\rangle] \geq E_{1^{\text{er}}\text{ état excité}} \quad \forall\ |\phi\rangle \perp |E_1\rangle \ , \quad (|E_1\rangle = \text{état fondamental exact}) \ . \quad (23.37)$$

On se trouve aussi dans cette situation quand le fondamental et le *premier* état excité sont de symétrie différente vis-à-vis d'une certaine opération ; c'est par exemple le cas si l'on sait que ce premier état a un spin $S(|E_2\rangle)$ non-nul alors que le fondamental est singulet. Alors, $|E_2\rangle$ désignant le premier état excité exact, le corollaire (23.27) garde sa validité :

$$S(|E_1\rangle) = 0,\ S(|E_2\rangle) > 0 \Longrightarrow E[|\phi\rangle \geq E_2 \quad \forall\ |\phi\rangle \perp |E_1\rangle \ , \quad (23.38)$$

$|\phi\rangle$ étant évidemment un état de spin (d'essai) ayant ayant un spin égal à $S(|E_2\rangle)$.

23.1.3 Méthode de Hartree - Fock

La méthode variationelle joue également un grand rôle pour la compréhension du comportement des assemblées de particules en interaction (Problème à N corps). Dans le cas des fermions en interaction, l'application la plus fameuse est probablement celle qui conduit aux équations de Hartree - Fock (voir par exemple [15] p. 662, [4] p. 535), qui joue un rôle de premier plan en Physique (nucléaire, atomique, moléculaire et du solide).

L'idée de base est celle de toutes les méthodes de champ moyen : on remplace le terme d'interaction à deux corps par un terme effectif représentant l'action moyenne de tous les électrons sur l'un d'entre eux. Autrement dit, on admet qu'il est possible

de trouver une approximation où finalement le terme à deux corps est remplacé par un terme ne dépendant que des variables d'un seul électron, l'action de tous les autres étant moyennée d'une certaine façon. Dès lors, on retrouve *formellement* une description où le Hamiltonien effectif est une somme de termes à un corps et possède donc des fonctions propres formées de produits de spin-orbitales. Compte tenu des nécessités imposées par le Principe de Pauli, la fonction générique de l'approximation de Hartree - Fock est donc un[10] déterminant de Slater, construit sur des spin-orbitales orthonormées :

$$\Psi_{\mathrm{A}}(1, 2, \ldots, N) = \sqrt{N!}\, Y_{\mathrm{A}} \prod_{i=1}^{N} \psi_{k_i}(i) \equiv \sqrt{N!}\, Y_{\mathrm{A}} \Phi(1, 2, \ldots, N) \ . \tag{23.39}$$

Par opposition avec un modèle à particules strictement indépendantes, où les spin-orbitales sont données d'avance par construction, les $\{\psi_{k_i}\}$ sont maintenant *inconnues* et toute la question est précisément de les trouver, c'est-à-dire d'écrire les équations qu'elles satisfont (équations[11] SCF), et de les résoudre.

En définitive, il s'agit de trouver la meilleure fonction du genre déterminant, puisque le déterminant est la forme de fonction retenant fondamentalement l'idée d'une théorie à *une* particule ; une méthode pour obtenir les équations SCF consiste à invoquer le principe variationnel, dont l'application va précisément conduire à la fonction optimisée. Il convient donc d'abord de calculer l'expression formelle de l'énergie moyenne E associée à un tel déterminant, puis d'effectuer la minimisation.

Pour N fermions de masse m en interaction, le Hamiltonien est[12] :

$$H = \sum_{i=1}^{N} \left[\frac{\vec{p}_i^{\,2}}{2m} + v(\vec{r}_i)\right] + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1,\, j\neq i}^{N} w(\vec{r}_i, \vec{r}_j) \equiv \sum_{i=1}^{N} \left[H_1(i) + W(\{\vec{r}_i\}_i)\right] \ . \tag{23.40}$$

Pour des électrons dans un atome de numéro atomique Z, et s'en tenant aux interactions de Coulomb, $v(r) = -\frac{Ze'^2}{r}$ et $w(\vec{r}_i, \vec{r}_j) = \frac{e'^2}{r_{ij}}$, où $r_{ij} = \|\vec{r}_i - \vec{r}_j\|$. Dans la suite, on suppose que l'interaction à deux corps ne dépend que des coordonnées, pas du spin des deux particules (l'abandon de cette hypothèse de travail ne pose aucune difficulté).

Le Hamiltonien (23.40) – comme toute observable relative à un ensemble de particules identiques –, commute avec toutes les permutations du groupe symétrique $\mathcal{S}_N$. En vertu de l'hermiticité et de l'idempotence de l'antisymétriseur Y_{A} (voir éq. (22.53)), et en prenant une fonction du type (23.39), on a :

$$E \equiv \langle \Psi_{\mathrm{A}} | H | \Psi_{\mathrm{A}} \rangle = N!\, \langle Y_{\mathrm{A}}\Phi | H | Y_{\mathrm{A}}\Phi \rangle = \langle \Phi | H | N! Y_{\mathrm{A}}\Phi \rangle \ . \tag{23.41}$$

H contient deux types de termes ; $H_1(i)$ n'agit que sur une particule à la fois ; en vertu de l'orthogonalité des spin-orbitales, tout terme du *ket* issu d'une simple transposition

[10] ou une combinaison linéaire de déterminants, selon la valeur du spin total S. Cette complication est ignorée dans la suite du formalisme, mais ne présente aucune difficulté.

[11] L'acronyme SCF est pour *self-consistent field* ; en français, on préférera *champ auto-cohérent* au barbarisme fréquent *champ self-consistent.*

[12] L'écriture à droite dans (23.40) ne doit pas dissimuler le fait que H est une somme d'opérateurs à *plusieurs* corps.

(la plus *petite* permutation) donnera zéro. Il reste donc :

$$E = \sum_{i=1}^{N} \langle \psi_{k_i} | H_1 | \psi_{k_i} \rangle \; + \; \sum_{\lambda=1}^{N!} (-1)^{\sigma_\lambda} \, \langle \Phi | W | P_\lambda \Phi \rangle \; . \tag{23.42}$$

Le terme W est une somme d'opérateurs w agissant chacun sur deux électrons, pas plus ; si donc la permutation P_λ implique plus de deux particules, le terme correspondant est nul, à nouveau par l'orthogonalité supposée des spin-orbitales. De ce fait, l'expression de l'énergie se réduit à :

$$E = \sum_{i=1}^{N} \langle \psi_{k_i} | H_1 | \psi_{k_i} \rangle \; + \; \frac{1}{2} \sum_{i \neq j} \langle \Phi | w \big(\mathbf{1} - \Sigma_{r,\, s} T_{rs} \big) | \Phi \rangle \tag{23.43}$$

où les T_{rs} sont les transpositions échangeant r et s (le signe $-$ vient de la signature des transpositions). Parmi tous les termes issus des deux sommes doubles, seuls ceux où i et j sont aussi les variables permutées donnent une contribution non-nulle. Au total, la valeur moyenne de l'énergie a pour expression :

$$E = \sum_{i=1}^{N} \langle \psi_{k_i} | H_1 | \psi_{k_i} \rangle + \frac{1}{2} \sum_{i \neq j} \left[\langle \psi_{k_i} \psi_{k_j} | w | \psi_{k_i} \psi_{k_j} \rangle - \langle \psi_{k_i} \psi_{k_j} | w | \psi_{k_j} \psi_{k_i} \rangle \right] \; , \tag{23.44}$$

soit :

$$\boxed{E = \sum_{i=1}^{N} \langle \psi_{k_i} | H_1 | \psi_{k_i} \rangle \; + \; \frac{1}{2} \sum_{i \neq j} \left(\mathcal{J}_{k_i k_j} \; - \; \mathcal{K}_{k_i k_j} \right)} \tag{23.45}$$

avec :

$$\mathcal{J}_{k_i k_j} \stackrel{\text{déf}}{=} \langle \psi_{k_i} \psi_{k_j} | w | \psi_{k_i} \psi_{k_j} \rangle \; , \qquad \mathcal{K}_{k_i k_j} \stackrel{\text{déf}}{=} \langle \psi_{k_i} \psi_{k_j} | w | \psi_{k_j} \psi_{k_i} \rangle \; . \tag{23.46}$$

Par convention, dans les éléments de matrice à deux corps, les variables muettes de sommation sont écrites de gauche à droite dans l'ordre naturel. Les éléments de matrice se calculent en effectuant les produits scalaires comme d'habitude : intégration sur les variables orbitales et sommation sur les variables de spin. La valeur moyenne de W apparaît ainsi comme une somme de deux termes appelés respectivement intégrales directes et d'échange et définis ici en termes de *spin-orbitales.*

Les équations SCF vont maintenant s'obtenir en effectuant une petite variation de tous les $\{\psi_{k_i}\}$ et en annulant la variation δE correspondante de l'énergie. La variation d'un *bra* $\langle \psi_{k_i} |$ donne simplement $\langle \delta\psi_{k_i} |$, celle de $\langle \psi_{k_i} \psi_{k_j} |$ donne $\langle (\delta\psi_{k_i}) \psi_{k_j} | + \langle \psi_{k_i} (\delta\psi_{k_j}) |$. Au total, on obtient la condition de stationnarité $\delta E = 0$ de la valeur moyenne E par rapport aux variations *indépendantes* de toutes les $\{\psi_{k_i}\}$ sous la forme d'une certaine somme ayant l'allure $\sum_i \langle \delta\psi_{k_i} | \mathcal{H}[\{\psi_{k_j}\}_j] \rangle - \varepsilon \sum_i \langle \delta\psi_{k_i} | \psi_{k_i} \rangle$+h.c. – typique d'une variation avec le paramètre de Lagrange ε –, où $|\mathcal{H}[\{\psi_{k_j}\}_j]\rangle$ désigne l'action d'un certain opérateur sur *l'ensemble* des fonctions inconnues $\{\psi_{k_j}\}_j$. Ainsi, partant de (23.45), la condition de stationnarité de E débouche à nouveau sur une équation aux valeurs propres, qui est non-linéaire puisque l'opérateur $\mathcal{H}$ *dépend* de ses fonctions propres. Explicitement,

les fonctions cherchées doivent satisfaire les équations suivantes[13], qui constituent les équations de Hartree - Fock :

$$\boxed{H_1(\vec{r}_1,\vec{p}_1)\psi_{k_i}(\vec{x}_1)+\sum_j\Big[\mathcal{W}[\psi_{k_j},\,\psi_{k_j}]\,\psi_{k_i}(\vec{x}_1)-\mathcal{W}[\psi_{k_j},\,\psi_{k_i}]\psi_{k_j}(\vec{x}_1)\Big]=\varepsilon\psi_{k_i}(\vec{x}_1)} \quad (23.47)$$

où les $\mathcal{W}$ sont des opérateurs définis comme :

$$\boxed{\mathcal{W}[\psi_{k_n},\,\psi_{k_{n'}}] \overset{\text{déf}}{=} \sum_{m_{s2}}\int \psi^*_{k_n}(\vec{x}_2)w(r_{12})\psi_{k_{n'}}(\vec{x}_2)\,\mathrm{d}^3r_2} \quad (23.48)$$

Les ψ_{k_i} et les valeurs propres ε sont les inconnues à déterminer (bien noter que ψ_{k_i} est en facteur de la première somme, mais est à *l'intérieur* de la seconde).

Ces équations sont relativement aisées à interpréter ; la première somme est :

$$\Big[\sum_{m_{s2}}\int\sum_j|\psi_{k_j}(\vec{x}_2)|^2\,w(r_{12})\,\mathrm{d}^3r_2\Big]\psi_{k_i}(\vec{x}_1)\ , \quad (23.49)$$

et n'est autre que l'interaction électrostatique entre un électron situé en $\vec{r}_1$ et la densité résultant de *tous* les autres, distribués suivant la densité $\sum_j|\psi_{k_j}(\vec{x}_2)|^2$: c'est ce qui affecte le mouvement d'un électron dans le champ électrostatique moyen de tous les autres. Ce terme est dit *de Hartree*, et serait seul présent si la fonction d'onde pouvait être choisie comme un simple produit (non-antisymétrisé) ; à lui seul, il justifirait déjà la terminologie de *champ moyen.* On définit parfois le Hamiltonien de Hartree $H_{\text{Hartree}}(\vec{r}_1,\vec{p}_1)$ comme :

$$H_{\text{Hartree}}(\vec{r}_1,\vec{p}_1) \overset{\text{déf}}{=} H_1(\vec{r}_1,\vec{p}_1)+\sum_{m_{s2}}\int\sum_j|\psi_{k_j}(\vec{r}_2)|^2\,w(r_{12})\,\mathrm{d}^3r_2\ ; \quad (23.50)$$

ceci permet de récrire les équations (23.47) comme :

$$H_{\text{Hartree}}(\vec{r}_1,\vec{p}_1)\psi_{k_i}(\vec{x}_1)-\sum_j\mathcal{W}[\psi_{k_j},\,\psi_{k_i}]\psi_{k_j}(\vec{x}_1)=\varepsilon\psi_{k_i}(\vec{x}_1)\ . \quad (23.51)$$

En revanche, la sommation complémentaire à H_{Hartree} dans (23.51), qui résulte directement des permutations à deux corps – c'est-à-dire du Principe de Pauli – n'a pas d'interprétation classique, ce qui n'est pas une surprise compte tenu de son origine.

En introduisant le Hamiltonien dit SCF, H_{SCF}, défini par son action sur une spin-orbitale comme reconstituant le premier membre de (23.47), cette dernière prend *formellement* l'allure d'une équation aux vecteurs propres :

$$H_{\text{SCF}}\big[\{|\psi_{k_j}\rangle\}_j\big]\psi_{k_i}=\varepsilon|\psi_{k_i}\rangle\ . \quad (23.52)$$

[13]Le facteur $\frac{1}{2}$, visible dans (23.45), disparaît dans la variation car les termes d'interaction sont quadratiques par rapport aux spin-orbitales (tout comme $\frac{1}{2}\delta(x^2)=x\delta x$).
Par ailleurs, on note $\vec{x}$ le couple $(\vec{r}, m_s)$.

Il faut cependant bien voir que l'opérateur au premier membre contient un terme intégral (non-local), et est une fonction explicite de ses vecteurs propres.

La résolution des équations de Hartree - Fock nécessite un ordinateur et procède par itération ; on se donne un jeu de fonctions au départ, avec lesquelles on calcule l'opérateur du premier membre, ou plus précisément sa matrice sur une base donnée. Après diagonalisation, on obtient un jeu de fonctions propres avec lesquelles le premier membre est recalculé, et ainsi de suite. Comme toute méthode itérative, le choix du point de départ est crucial pour la rapidité de la convergence[14]. Une fois celle-ci obtenue, on dispose donc d'un jeu de spin-orbitales dites *auto-cohérentes*, satisfaisant :

$$\boxed{H_{\text{SCF}}\big[\{|\psi_\lambda^{\text{SCF}}\rangle\}_\lambda\big]\psi_\lambda^{\text{SCF}} = \varepsilon_\lambda^{\text{SCF}}|\psi_\lambda^{\text{SCF}}\rangle} \tag{23.53}$$

L'énergie associée au déterminant Ψ_{SCF} construit sur ces solutions n'est pas égale à $\sum_\lambda \varepsilon_\lambda^{\text{SCF}}$ – alors on prendrait en compte deux fois l'interaction au sein d'une même paire –, mais doit être calculée en prenant la valeur moyenne du Hamiltonien de départ avec Ψ_{SCF} :

$$E_{\text{SCF}} = \langle\Psi_{\text{SCF}}|H|\Psi_{\text{SCF}}\rangle \ , \tag{23.54}$$

ce qui revient à retrancher *une fois* de la somme des valeurs propres SCF la valeur moyenne du terme d'interaction $W(\{r_{ij}\}_{i,j})$.

Bien évidemment, on a :

$$E_{\text{SCF}} \geq E_{\text{exacte}} \ ; \tag{23.55}$$

de façon (très) conventionnelle, on appelle *énergie de corrélation* E_{corr} la différence entre ces deux énergies :

$$E_{\text{corr}} \stackrel{\text{déf}}{=} E_{\text{SCF}} - E_{\text{exacte}} \geq 0 \ ; \tag{23.56}$$

la terminologie est peu heureuse : au sein d'un schéma de Hartree - Fock, les électrons sont *déjà* corrélés – indépendamment de la corrélation indissolublement liée au Principe de Pauli déjà présente en l'absence de toute interaction au sens classique du terme –, puisque leur interaction $w(r_{ij})$ est *partiellement* prise en compte par le traitement de champ moyen.

Le plus souvent, on écrit la valeur moyenne du terme d'interaction dans (23.45) à l'aide d'intégrales directes et d'échange définies avec des *orbitales* ; la correspondance se trouve comme suit. Si on écrit chaque spin-orbitale sous la forme :

$$\psi_{k_i}(\vec{x}_1) = \phi_{k_i}(\vec{r}_1)\,\xi_i(m_{s1}) \ , \tag{23.57}$$

alors :

$$\mathcal{J}_{k_ik_j} = \int \phi_{k_i}^*(\vec{r}_1)\phi_{k_j}^*(\vec{r}_2)w(r_{12})\phi_{k_i}(\vec{r}_1)\phi_{k_j}(\vec{r}_2)\,\mathrm{d}^3r_1\mathrm{d}^3r_2 \ \times \sum_{m_{s1}}\sum_{m_{s2}} \xi_i(m_{s1})\xi_i(m_{s1})\xi_j(m_{s2})\xi_j(m_{s2}) \ ; \tag{23.58}$$

[14]La convergence est parfois problématique, et peut être sujette à des instabilités, ou à des brisures spontanées de symétrie vis-à-vis du spin total.

la somme sur les spins reconstitue $\langle\xi_i|\xi_i\rangle\langle\xi_j|\xi_j\rangle$, qui vaut 1, d'où :

$$\mathcal{J}_{k_ik_j} = \int \phi^*_{k_i}(\vec{r}_1)\phi^*_{k_j}(\vec{r}_2)w(r_{12})\phi_{k_i}(\vec{r}_1)\phi_{k_j}(\vec{r}_2)\,\mathrm{d}^3r_1\mathrm{d}^3r_2 \stackrel{\text{déf}}{=} J_{ij} \ . \tag{23.59}$$

De même, pour les intégrales d'échange, on a :

$$\mathcal{K}_{k_ik_j} = \int \phi^*_{k_i}(\vec{r}_1)\phi^*_{k_j}(\vec{r}_2)w(r_{12})\phi_{k_j}(\vec{r}_1)\phi_{k_i}(\vec{r}_2)\,\mathrm{d}^3r_1\mathrm{d}^3r_2 \ \times \sum_{m_{s1}}\sum_{m_{s2}}\xi_j(m_{s1})\xi_i(m_{s1})\xi_i(m_{s2})\xi_j(m_{s2}) \ ;$$

cette fois, la somme sur les spins est $\langle\xi_j|\xi_i\rangle\langle\xi_i|\xi_j\rangle$, qui vaut 0 ou 1, – et est, de ce fait, égal à son propre carré – d'où :

$$\mathcal{K}_{k_ik_j} = \int \phi^*_{k_i}(\vec{r}_1)\phi^*_{k_j}(\vec{r}_2)w(r_{12})\phi_{k_j}(\vec{r}_1)\phi_{k_i}(\vec{r}_2)\,\mathrm{d}^3r_1\mathrm{d}^3r_2\,\langle\xi_j|\xi_i\rangle \stackrel{\text{déf}}{=} K_{ij}\,\langle\xi_j|\xi_i\rangle \ , \tag{23.60}$$

Avec ces définitions, les intégrales directes J_{ij} et d'échange K_{ij} sont relatives aux *orbitales.* Remarquons aussi que dans la somme double au second membre de (23.45), on peut relâcher la contrainte $i \neq j$ puisque le terme correspondant est nul.

♦ *Remarques*

1. Dans l'immense majorité des cas, les opérateurs se présentent sous la forme d'une somme de termes à deux corps au plus :

$$\Omega(\vec{x}_1, \vec{x}_2, ..., \vec{x}_N) = \sum_{i=1}^{N}\omega_1(x_i) + \frac{1}{2}\sum_{i\neq j}\omega_2(x_i,\, x_j) \equiv \Omega_1 + \Omega_2 \ , \tag{23.61}$$

Pour N particules identiques, la fonction d'onde doit être symétrique ou antisymétrique. Compte tenu de cette propriété, toutes les valeurs moyennes peuvent s'exprimer à l'aide de deux opérateurs densité réduits à une et deux particules, définis comme suit (voir ausssi problème 22.6.4) :

$$\gamma(\vec{x}_1\,;\,\vec{x}\,'_1) \stackrel{\text{déf}}{=} \int \Psi(\vec{x}_1, \vec{x}_2, ..., \vec{x}_N)\Psi^*(\vec{x}\,'_1, \vec{x}_2, ..., \vec{x}_N)\,\mathrm{d}x_2\mathrm{d}x_3...\mathrm{d}x_N \ , \tag{23.62}$$

$$\Gamma(\vec{x}_1,\,\vec{x}_2\,;\,\vec{x}\,'_1,\,\vec{x}\,'_2) \stackrel{\text{déf}}{=} \int \Psi(\vec{x}_1, \vec{x}_2, ..., \vec{x}_N)\Psi^*(\vec{x}\,'_1, \vec{x}\,'_2, ..., \vec{x}_N)\,\mathrm{d}x_3\mathrm{d}x_4...\mathrm{d}x_N \ . \tag{23.63}$$

En effet, en écrivant l'expression des valeurs moyennes, et en jouant avec la symétrie vis-à-vis des variables muettes, on trouve aisément[15] :

$$\langle\Omega_1\rangle = N\int \mathrm{d}x_1\{\omega(x_1)\gamma(\vec{x}_1\,;\,\vec{x}\,'_1)\}_{\vec{x}\,'_1=\vec{x}_1} \equiv N\mathrm{Tr}(\Omega_1\gamma) \ , \tag{23.64}$$

[15]Les deux égalités suivantes se comprennent bien : en raison de l'identité des particules, celles-ci jouent toutes le même rôle ; la valeur moyenne d'une somme d'opérateurs à un corps est simplement N fois la moyenne pour l'une des particules ; celle d'une somme d'opérateurs à deux corps vaut celle d'une paire, multipliée par le nombre de paires distinctes, soit $\frac{1}{2}N(N-1)$.

$$\langle \Omega_2 \rangle = \frac{1}{2}N(N-1)\int \mathrm{d}x_1\mathrm{d}x_2\{\omega(x_1,\, x_2)\Gamma(\vec{x}_1,\, \vec{x}_2\,;\, \vec{x}'_1,\, \vec{x}'_2)\}_{\vec{x}'_1=\vec{x}_1,\, \vec{x}'_2=\vec{x}_2} \equiv \frac{1}{2}N(N-1)\mathrm{Tr}(\Omega_2\Gamma)\ . \quad (23.65)$$

Autrement dit, la fonction d'onde contenant finalement trop d'information – en raison de son obligatoire symétrie de permutation –, il est suffisant d'obtenir les deux opérateurs densité réduits, pour lesquels on peut écrire les équations SCF, et les résoudre par un procédé itératif.

Pour les fermions, et avec l'apport du théorème de Hohenberg et Kohn [176], ces éléments sont le point de départ des méthodes très variées dites *de fonctionnelle de la densité*. Ce théorème affirme d'une part que l'énergie (fondamentale) E d'un système de fermions en interaction ne dépend que de la densité à un corps $n(\vec{x}) \stackrel{\text{déf}}{=} N\gamma(\vec{x},\, \vec{x})$ dans l'état fondamental : E est une certaine *fonctionnelle* de $n(\vec{x})$, $E[n(\vec{x})]$; d'autre part, cette fonctionnelle est *minimale* si la densité est celle de l'état fondamental. Toute la question est évidemment d'obtenir la fonctionnelle inconnue $E[n(\vec{x})]$, que l'on ne sait pas trouver exactement, d'où les nombreuses méthodes d'approximation.

2. L'*ansatz* de Bethe n'est pas à proprement parler une méthode d'approximation puisqu'il permet d'obtenir la solution *exacte* de certains modèles, bien particuliers, de particules en interaction (voir par exemple le cas de la chaîne de spins $S = 1/2$, chapitre 30, sous-section 30.6.2). La mise en œuvre, le plus souvent fort technique, est exposée en détails dans le livre de Gaudin [177]. ♦

23.2 Théorie des perturbations stationnaires

La théorie des perturbations stationnaires est une autre méthode de résolution approchée de l'équation aux valeurs et vecteurs propres :

$$H|\psi\rangle = E|\psi\rangle\ . \quad (23.66)$$

Cette approche peut être envisagée à chaque fois que le Hamiltonien peut être décomposé en deux morceaux comme suit :

$$H = H_0 + gV \quad (23.67)$$

où H_0 est un Hamiltonien dont les solutions exactes sont connues, appelé Hamiltonien non-perturbé. V est un opérateur hermitique (ici indépendant du temps) homogène à une énergie ; g est un nombre réel sans dimension introduit pour la commodité et constituant également une mesure du terme additionnel gV, appelé perturbation[16]. Au total, pour que la méthode fonctionne bien, il faut que gV soit *petit*, dans un sens à préciser autant

[16] g est aussi là pour compter facilement les puissances. À la fin, on fera formellement $g = 1$ pour se recaler sur l'écriture usuelle : il faudra alors comprendre que se tient caché dans V un petit paramètre multiplicatif qui autorise (sans le justifier complètement) l'usage d'une théorie de perturbation.

que faire se peut[17]. Par exemple, H_0 est le Hamiltonien d'un atome isolé et gV décrit l'interaction de celui-ci avec un champ extérieur, ici supposé statique, de petite intensité ; en pareil cas, g contiendra typiquement le module du champ extérieur – plus précisément, g sera le rapport entre l'énergie typique de couplage avec le champ et une énergie relative au système non-perturbé. Pour un Hamiltonien de la forme $H = H_0 + V$, il suffit de faire $g = 1$ dans toutes les formules ci-dessous.

La théorie des perturbations joue un rôle majeur en Physique quantique : compte tenu de la relative complexité du formalisme, la classe de problèmes que l'on sait résoudre exactement est très réduite. Une telle situation n'est d'ailleurs pas propre au domaine quantique : le problème des trois corps en interaction gravitationnelle – pour lequel ont d'ailleurs été développées des méthodes de type perturbatif – est d'une redoutable complexité[18]. En outre, les méthodes de perturbation s'inscrivent très naturellement dans la démarche des physiciens : après avoir *dégrossi* un problème donné en retenant les phénomènes dominants (représentés par H_0), on affine la description en incluant de "petits" effets qui, malgré leur petitesse, peuvent jouer un rôle important, parfois spectaculaire (éclatement des raies spectrales, réduction de la durée de vie d'un niveau excité par contamination, etc.). D'une façon générale, la Théorie des perturbations joue un rôle de premier plan dans le problème à N-corps, pour lequel il a d'ailleurs fallu mettre au point des méthodes plus raffinées pour rendre compte d'effets collectifs[19].

L'idée fondamentale sur laquelle repose la méthode des perturbations est qu'il existe un développement en puissances *entières* de la constante de coupage g pour les modes propres de H ; autrement dit, on admet que toutes les quantités d'intérêt sont des fonctions *analytiques* de g dans un voisinage de l'origine. Dans toute la suite, on suppose donc que tout état propre $|\psi\rangle$ de H et la valeur propre E associée admettent

[17] La *petitesse* est visiblement une condition *nécessaire* ; formellement, elle n'est pas en général suffisante – voir les remarques ultérieures à ce sujet.

[18] En 1912, Sundman [178] a démontré l'existence d'une solution analytique du problème des trois corps en interaction gravitationnelle sous la forme d'une série en $t^{1/3}$. Toute la question est alors de déterminer le rayon de convergence de cette série, qui est fixé par les singularités que constituent les *collisions* à deux et trois corps. L'un des éléments forts de la démonstration consiste à préciser les conditions sur l'état initial écartant toute possibilité de collision triple ; cela étant fait, il est possible de préciser le rayon de convergence de la série qui, malheureusement, est si lentement convergente qu'elle est, en pratique, quasiment inutilisable (sur l'histoire du théorème de Sundman et son contenu précis, voir l'article récent de Henkel [179]).

Le roi Oscar II de Suède, sur le conseil de Gösta Mittag - Leffler, a créé un prix destiné à récompenser qui trouverait la solution du problème à N-corps :

> *"Given a system of arbitrarily many mass points that attract each according to Newton's law, under the assumption that no two points ever collide, try to find a representation of the coordinates of each point as a series in a variable that is some known function of time and for all of whose values the series converges uniformly."*

Le prix a finalement été décerné à Poincaré (1900) pour ses contributions majeures en la matière – qui ont ouvert la voie à la théorie du chaos –, en dépit du fait que le problème n'avait pas été résolu.

[19] Essentiellement, il s'agit de méthodes qui mettent le doigt sur les termes les plus importants et qui, moyennant des resommations à tous les ordres (à l'aide de la résolvante par exemple) d'une *classe* de processus élémentaires, dépassent et de loin la technique élémentaire exposée en détail dans ce chapitre. Sur ces questions, deux ouvrages sont particulièrement recommandés, ceux de March et al. [180] et de Fetter et Walecka [174], où il est fait un usage intensif des diagrammes de Feynman. Mattuck [181] donne de ces derniers une présentation très pédagogique.

des développements du genre :

$$|\psi\rangle = |\psi^{(0)}\rangle + g|\phi_1\rangle + g^2|\phi_2\rangle + \dots \ , \tag{23.68}$$

$$E = E^{(0)} + g\varepsilon_1 + g^2\varepsilon_2 + \dots \ . \tag{23.69}$$

Les $\{|\phi_n\rangle\}$ et les $\{\varepsilon_n\}$, $n \geq 1$ constituent les corrections successives à l'ordre n pour le vecteur et l'énergie propres. Une telle hypothèse est loin d'être assurée : l'obligation pour tout mode propre de H de se confondre avec ceux de H_0 pour $g = 0$ peut être satisfaite de bien d'autres manières. Par exemple, l'énergie E pourrait avoir un développement du genre :

$$E = E^{(0)}(1 + g \ln g + \dots) \ , \tag{23.70}$$

qui se réduit bien à $E^{(0)}$ pour $g = 0$ mais qui n'est pas pour autant développable en série entière au voisinage de $g = 0$: si on essaie formellement un développement de Taylor, on voit que chaque coefficient du développement est infini... Autre exemple, soit :

$$E = E^{(0)}(1 + \mathrm{e}^{-1/g} + \dots) \tag{23.71}$$

une telle expression a un développement en série entière ... dont tous les coefficients sont cette fois nuls, traduisant le fait que $\mathrm{e}^{-1/g}$ tend vers zéro plus vite que toute puissance finie de g. De tels développements sont loin d'être rares et surviennent, par définition, pour les perturbations dites *singulières*, qui exigent la mise en œuvre d'une méthode particulière, éventuellement spécifique au cas traité[20] ; il sera clair que la méthode exposée ci-dessous est totalement inopérante en pareille situation, puisque les formules obtenues plus loin donneraient alors des termes tous infinis, ou tous nuls. Dans le cas contraire où les développements en puissances (23.68) et (23.69) existent, la perturbation est dite régulière ; seules les perturbations de ce dernier type sont considérées dans la suite.

En outre, même quand un développement en puissances entières existe ou semble exister – au sens où les premiers coefficients sont finis – il conviendrait d'examiner la convergence de la série qui s'amorce[21]. On découvrirait alors qu'il s'agit souvent en fait d'un développement *asymptotique* [183], fort utile, mais qu'il convient de manier avec précaution[22]. Ces subtilités seront écartées dans la suite, par pragmatisme, mais

[20]Un exemple célèbre est la théorie standard de la supraconductivité (théorie BCS), qui repose sur la notion de *paires de Cooper*. Comme on le verra, l'énergie de liaison d'une paire de Cooper est de la forme e^{-A/V_0}, où $V_0 > 0$ est l'interaction entre les électrons de la paire (voir chapitre 32, section 32.5). Un tel résultat ne peut être obtenu par perturbation, même si (et c'est le cas) l'interaction V_0 est très petite.

[21]Cette question est, dans la quasi-totalité des cas, totalement académique. Examiner la convergence suppose que l'on connaît l'expression générale du terme d'ordre n, ce qui n'est vérifié en pratique que pour des cas presque triviaux.

[22]Un exemple aide à comprendre ce point. Soit la fonction $I(g)$ définie, pour $g > 0$, par :

$$I(g) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{\mathrm{e}^{-x/g}}{1+x}\, \mathrm{d}x \ . \tag{23.72}$$

De toute évidence, $\lim_{g \to 0_+} I(g) = 0$. L'exponentielle agit comme une fonction de coupure et l'intégrand est quasiment nul dès que x est à peine supérieur à g ; donc quand $0 < g \ll 1$, seul le voisinage de l'origine contribue notablement à l'intégrale. On peut alors tenter une opération illicite, consistant à développer $(1 + x)^{-1}$ en série entière, bien que ce développement ne converge que pour $|x| < 1$ et alors que l'intervalle d'intégration va jusqu'à $+\infty$. En poussant l'audace jusqu'à échanger sommation et

il convient de les avoir présentes à l'esprit. Quoi qu'il en soit, le physicien est supposé avoir assez de flair pour se convaincre d'avance de la validité de la méthode qu'il choisit au cas par cas (voir par exemple la remarque p. 1074 ci-dessous concernant l'effet Stark).

Le chapitre VIII de la *Mécanique céleste* de Poincaré [182] commence par le commentaire suivant[a] :

> *"Il y a entre les géomètres et les astronomes une sorte de malentendu au sujet de la signification du mot convergence. Les géomètres préoccupés de la parfaite rigueur et souvent trop indifférents à la longueur de calculs inextricables dont ils conçoivent la possibilité, sans songer à les entreprendre effectivement, disent qu'une série est convergente quand la somme des termes tend vers une limite déterminée, quand même les premiers termes diminueraient très lentement. Les astronomes, au contraire, ont coutume de dire qu'une série converge quand les 20 premiers termes, par exemple, diminuent très rapidement, quand même les termes suivants devraient croître indéfiniment. Ainsi, pour prendre un exemple simple, considérons les deux séries qui ont pour terme général :*
>
> $$\frac{1\,000^n}{n!} \qquad \frac{n!}{1\,000^n}$$
>
> *Les géomètres diront que la première converge, et même qu'elle converge rapidement, parce que le millionième terme est beaucoup plus petit que le 999 999ème ; mais ils regarderont la seconde comme divergente, parce que le terme général peut croître au-delà de toute limite. Les astronomes, au contraire, regarderont la première série comme divergente, parce que les 1 000 premiers termes vont en croissant ; et la seconde comme convergente, parce que les 1 000 premiers termes vont en décroissant et que cette décroissance est d'abord très rapide. Les deux règles sont légitimes : la première, dans les recherches théoriques ; la seconde, dans les applications numériques. Toutes deux doivent régner, mais dans deux domaines séparés et dont il importe de bien connaître les frontières".*

[a]Cité en `http://fr.wikipedia.org/wiki/Théorie_des_perturbations`.

"The series is divergent ; therefore we may be able to do something with it"
(Oliver Heaviside)

intégration, on obtient :

$$I(g) = \sum_{n=0}^{+\infty}(-1)^n \int_0^{+\infty} x^n \mathrm{e}^{-x/g}\,\mathrm{d}x = \sum_{n=0}^{+\infty}(-1)^n n!\, g^{n+1} \ . \tag{23.73}$$

La série de droite est visiblement divergente (mais vaut bien zéro en $g = 0$), son rayon de convergence est nul. Toutefois, pour $0 < g \ll 1$, le développement (23.73) judicieusement tronqué à ses N premiers termes (N est d'une certaine façon fonction de la valeur de g !) donne une approximation remarquable de la valeur prise par la fonction $I(g)$.

L'absence d'un développement en série entière pour $I(g)$ est d'ailleurs évident quand on observe que $I(0) = 0$ et que si on fait brutalement $g = 0_-$ dans (23.72), on trouve $I(0_-) = +\infty \ldots$: $g = 0$ est donc un point singulier de la fonction I – que l'on connaît d'ailleurs en tant que fonction spéciale bien répertoriée (I s'exprime à l'aide de $\mathrm{Ei}(x)$, fonction appelée exponentielle intégrale). Cette fonction a un comportement logarithmique près de $g = 0$ et on peut la prolonger dans le demi-plan complexe $\Re g < 0$, coupé du demi-axe réel négatif ; ce prolongement est possible : il suffit d'incliner continûment la demi-droite d'intégration en rapport avec l'argument de g de sorte qu'à l'infini l'argument du produit gx soit positif ; l'origine est un point de singularité en ce sens qu'il s'agit d'un point de branchement.

La technique exposée dans la suite est plus précisément la théorie de perturbations dite *de Rayleigh - Schrödinger*, sans conteste la plus utilisée. Une autre méthode existe, due à Brillouin et Wigner (pour une introduction, voir [180], et problème 23.4.3) où, essentiellement, la valeur propre E ne fait pas l'objet d'un développement en série, soit parce qu'elle est connue par d'autres moyens, soit parce qu'elle est déterminée en cours de route par la résolution d'une équation implicite.

Dans le but de sérier les difficultés, il est préférable de traiter séparément le cas de la perturbation d'un niveau non-dégénéré et celui où existe une dégénérescence. De surcroît, afin de s'en tenir à une théorie élémentaire, on suppose le spectre de H_0 entièrement discret ; il en résulte que, par hypothèse, les $\{|\psi_n^{(0)}\rangle\}$ forment une base complète[23].

23.2.1 Cas d'un niveau non-dégénéré

On suppose donc connues toutes les solutions propres de H_0 :

$$H_0|\psi_m^{(0)}\rangle = E_m^{(0)}|\psi_m^{(0)}\rangle \ . \qquad (23.74)$$

Les $|\psi_m^{(0)}\rangle$ sont normalisés à l'unité, et on se propose de calculer l'effet de gV sur un état propre donné, choisi une fois pour toutes, soit $|\psi_n^{(0)}\rangle$; par hypothèse, l'énergie $E_n^{(0)}$ est ici non-dégénérée. Il s'agit alors de résoudre :

$$(H_0 + gV)|\psi_n\rangle = E_n|\psi_n\rangle \qquad (23.75)$$

où E_n et $|\psi_n\rangle$ notent les énergie et vecteur propre qui tendent vers $E_n^{(0)}$ et $|\psi_n^{(0)}\rangle$ lorsque g tend vers zéro. La méhode de Rayleigh - Schrödinger repose sur l'hypothèse de l'existence des développements suivants[24] :

$$\boxed{|\psi_n\rangle = |\psi_n^{(0)}\rangle + g|\phi_{n1}\rangle + g^2|\phi_{n2}\rangle + \ldots \equiv |\psi_n^{(0)}\rangle + |\phi_n\rangle} \qquad (23.76)$$

$$\boxed{E_n = E_n^{(0)} + g\varepsilon_{n1} + g^2\varepsilon_{n2} + \ldots \equiv E_n^{(0)} + \varepsilon_n} \qquad (23.77)$$

En reportant ces développements dans (23.75), et en identifiant les termes de même puissance de g, on obtient la suite d'équations :

$$H_0|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle \ , \qquad (23.78)$$

[23]Pour tout dire, cette restriction est de pure forme. En pratique, cela n'empêchera nullement d'utiliser la théorie des perturbations dans des cas où l'on sait pertinemment que cette hypothèse n'est pas vérifiée ; le cas typique est celui d'un atome, dont le spectre n'est jamais entièrement discret (voir le problème 23.4.10 pour un cas d'école).

On appliquera aussi cette théorie pour la description de l'effet Stark sur l'hydrogène ... dont le spectre, en présence du champ électrique, n'a plus *aucun* état lié : dès qu'un champ électrique est présent, la probabilité d'ionisation n'est pas nulle, aussi petit que soit le champ, et tous les niveaux (discrets) acquièrent une durée de vie *finie*. Cependant, le passage par effet-tunnel est si faible à petit champ que le négliger est tout à fait acceptable, ce qui revient à continuer à raisonner en termes d'états liés. Comme toujours, les subtilités mathématiques doivent être appréciées au regard d'arguments physiques raisonnés.

[24]C'est à ce point que la méthode se sépare de celle de Brillouin - Wigner, où l'énergie E_n est gardée telle quelle, dans l'idée de savoir résoudre une équation implicite la déterminant (voir problème 23.4.3).

$$H_0|\phi_{n1}\rangle + V|\psi_n^{(0)}\rangle = E_n^{(0)}|\phi_{n1}\rangle + \varepsilon_{n1}|\psi_n^{(0)}\rangle \ , \tag{23.79}$$

$$H_0|\phi_{n2}\rangle + V|\phi_{n1}\rangle = E_n^{(0)}|\phi_{n2}\rangle + \varepsilon_{n1}|\phi_{n1}\rangle + \varepsilon_{n2}|\psi_n^{(0)}\rangle \ , \qquad \text{etc.} \tag{23.80}$$

Ces équations se récrivent[25], omettant la première en raison de sa trivialité :

$$(H_0 - E_n^{(0)})|\phi_{n1}\rangle = (\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle \ , \tag{23.82}$$

$$(H_0 - E_n^{(0)})|\phi_{n2}\rangle = (\varepsilon_{n1} - V)|\phi_{n1}\rangle + \varepsilon_{n2}|\psi_n^{(0)}\rangle \ , \quad \text{etc.} \tag{23.83}$$

Toutes ces équations ont la particularité d'avoir au premier membre la correction d'ordre $k+1$ à l'état propre, exprimée en termes des corrections d'ordre $\leq k$, présentes au second membre : par rapport à l'inconnue $|\phi_{nk+1}\rangle$, chaque équation est *inhomogène*. Cette série d'équations permet en principe de calculer de proche en proche toutes les corrections les unes après les autres : si on arrive à trouver ε_{n1}, (23.82) donne la première correction $|\phi_{n1}\rangle$, laquelle permettra de trouver ε_{n2} et, par l'équation suivante de trouver $|\phi_{n2}\rangle$, et ainsi de suite.

À ce stade, une remarque permet de simplifier un peu l'algèbre. Supposons le problème résolu ; on va notamment trouver les $|\phi_{nk}\rangle$ exprimés sur la base propre $\{|\psi_m^{(0)}\rangle\}_m$ et ayant en particulier *a priori* une composante non nulle sur l'état $|\psi_n^{(0)}\rangle$ dont on cherche les corrections ; cette dernière peut être exhibée en écrivant les choses comme suit :

$$|\phi_{nk}\rangle = c_{nnk}|\psi_n^{(0)}\rangle + \sum_{m\neq n} c_{nmk}|\psi_m^{(0)}\rangle \ ; \ldots \tag{23.84}$$

reportant cette expression dans le développement (23.68) et rassemblant tous les termes qui sont une composante le long de $|\psi_n^{(0)}\rangle$, on trouve :

$$|\psi_n\rangle = (1 + gc_{nn1} + g^2c_{nn2} + \ldots)|\psi_n^{(0)}\rangle + \text{ termes en } |\phi_{n\,k\geq 1}\rangle \ . \tag{23.85}$$

Compte tenu du regroupement qui vient d'être effectué, tous les termes en $|\phi_{n\,k\geq 1}\rangle$ ne contiennent plus $|\psi_n^{(0)}\rangle$ et sont donc tous orthogonaux à $|\psi_n^{(0)}\rangle$. Par ailleurs, comme toute équation aux valeurs et vecteurs propres, l'équation que l'on cherche à résoudre (approximativement) est une équation homogène ; toutes ses solutions sont donc au mieux définies à un facteur global près, qui doit être déterminé en temps utile sur d'autres considérations (par exemple, normalisation à l'unité). Ceci permet de *convenir* que la composante de $|\psi_n\rangle$ sur l'état non-perturbé $|\psi_n^{(0)}\rangle$ est égale à l'unité ; cette convention étant définitivement adoptée, la correction globale de $|\psi_n\rangle$ est orthogonale à $|\psi_n^{(0)}\rangle$:

$$\boxed{|\psi_n\rangle = |\psi_n^{(0)}\rangle + \text{ corrections orthogonales à } |\psi_n^{(0)}\rangle} \tag{23.86}$$

entraînant, à tous les ordres :

$$\boxed{|\phi_{nk}\rangle \perp |\psi_n^{(0)}\rangle} \tag{23.87}$$

[25]Pour simplifier les écritures, on omet l'opérateur identité $\mathbf{1}$ en facteur des scalaires :

$$H_0 - E_n^{(0)} \equiv H_0 - E_n^{(0)}\mathbf{1} \ , \qquad \varepsilon_{n1} - V \equiv \varepsilon_{n1}\mathbf{1} - V, \quad \text{etc.} \tag{23.81}$$

Examinons maintenant l'équation (23.82), dont l'inversion doit être conduite avec précaution ; en effet, la relation $(H_0 - E_n^{(0)})|\psi_n^{(0)}\rangle = 0$ montre que l'opérateur $(H_0 - E_n^{(0)})$ est un opérateur singulier, qui n'a pas d'inverse dans *tout* l'espace des états. Compte tenu de la propriété d'orthogonalité (23.87), une méthode de partition utilisant les projecteurs s'avère très utile ; posons :

$$P_n = |\psi_n^{(0)}\rangle\,\langle\psi_n^{(0)}| \ , \qquad Q_n = \mathbf{1} - P_n \equiv \sum_{m\neq n} |\psi_m^{(0)}\rangle\,\langle\psi_m^{(0)}| \ . \tag{23.88}$$

Comme les $\{|\psi_m^{(0)}\rangle\}$ sont orthogonaux et normalisés à l'unité, P_n est un projecteur :

$$P_n^2 = P_n \ , \qquad P_n^\dagger = P_n, \tag{23.89}$$

et il en va évidemment de même pour Q_n ; en outre $P_nQ_n = Q_nP_n = 0$. Par ailleurs :

$$P_nH_0 = |\psi_n^{(0)}\rangle\,\langle\psi_n^{(0)}|H_0 = |\psi_n^{(0)}\rangle\,\langle\psi_n^{(0)}|E_n^{(0)} = E_n^{(0)}P_n \ ; \tag{23.90}$$

et, par l'orthogonalité (23.87) :

$$P_n|\phi_{n1}\rangle = 0 \ . \tag{23.91}$$

Multipliant maintenant à gauche par P_n l'équation (23.82), il vient :

$$P_n(H_0 - E_n^{(0)})|\phi_{n1}\rangle = P_n(\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle \ . \tag{23.92}$$

Le premier membre est nul en vertu de (23.90) et de (23.91), d'où $P_n(\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle = 0$; multipliant par $\langle\psi_n^{(0)}|$ à gauche, il vient $\langle\psi_n^{(0)}|(\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle = 0$, soit :

$$\boxed{\varepsilon_{n1} = \langle\psi_n^{(0)}|V|\psi_n^{(0)}\rangle} \tag{23.93}$$

Ceci constitue le premier résultat important, qui s'énonce :

> *Pour un niveau non-dégénéré, la première correction ε_{n1} à l'énergie est égale à la valeur moyenne de la perturbation dans l'état non-perturbé.*

Cette affirmation semble après coup presque évidente. En faisant formellement $g = 1$, donc avec $H = H_0 + V$, l'énergie de l'état $|\psi_n\rangle$ est :

$$\boxed{E_n = E_n^{(0)} + \langle\psi_n^{(0)}|V|\psi_n^{(0)}\rangle + \text{termes d'ordre supérieur}} \tag{23.94}$$

Bien sûr, il peut arriver que la moyenne de V dans l'état non perturbé $|\psi_n^{(0)}\rangle$ soit nulle (pour des raisons de symétrie, par exemple) ; en pareil cas, le *shift* en énergie est au moins du second ordre. En tout état de cause, la première correction formelle à l'énergie ne dépend que de l'état propre à l'ordre zéro.

Une fois obtenue la première correction pour l'énergie, on peut trouver la première correction à l'état propre. Revenant à (23.82) et en multipliant cette fois à gauche par Q_n, on obtient :

$$Q_n(H_0 - E_n^{(0)})|\phi_{n1}\rangle = Q_n(\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle \ . \tag{23.95}$$

Maintenant on peut insérer $\mathbf{1} = P_n + Q_n$ au premier membre, juste devant $|\phi_{n1}\rangle$. Comme H_0 est diagonal, $Q_n H_0 P_n = 0$ et il reste :

$$Q_n(H_0 - E_n^{(0)})Q_n|\phi_{n1}\rangle = Q_n(\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle \ . \tag{23.96}$$

$Q_n|\psi_n^{(0)}\rangle$ étant nul, la même insertion au second membre, juste devant $|\psi_n^{(0)}\rangle$, donne :

$$Q_n(H_0 - E_n^{(0)})Q_n|\phi_{n1}\rangle = Q_n\,(\varepsilon_{n1} - V)\,P_n|\psi_n^{(0)}\rangle = -\,Q_n\,V|\psi_n^{(0)}\rangle \tag{23.97}$$

(on peut omettre P_n devant $|\psi_n^{(0)}\rangle$ puisque $P_n|\psi_n^{(0)}\rangle = |\psi_n^{(0)}\rangle$). $Q_n(H_0 - E_n^{(0)})Q_n$ est la projection de $H_0 - E_n^{(0)}$ dans le sous-espace orthogonal à $|\psi_n^{(0)}\rangle$. Ceci peut maintenant s'inverser : puisque l'énergie $E_n^{(0)}$ est non-dégénérée, il n'y a aucune valeur propre égale à $E_n^{(0)}$ dans le sous-espace sur lequel Q_n projette, de sorte que l'opérateur au premier membre de (23.97) n'a aucune valeur propre nulle et possède un inverse, d'où :

$$\boxed{|\phi_{n1}\rangle = -\,[Q_n\,(H_0 - E_n^{(0)})\,Q_n]^{-1}\,V|\psi_n^{(0)}\rangle} \tag{23.98}$$

Par définition, on a :

$$[Q_n\,(H_0 - E_n^{(0)})\,Q_n]^{-1} = \sum_{m\neq n} |\psi_m^{(0)}\rangle\,\frac{1}{E_m^{(0)} - E_n^{(0)}}\,\langle\psi_m^{(0)}| \ , \tag{23.99}$$

ce qui donne l'expression explicite :

$$|\phi_{n1}\rangle = \sum_{m\neq n} \frac{\langle\psi_m^{(0)}|V|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}\,|\psi_m^{(0)}\rangle \ . \tag{23.100}$$

Faisant maintenant $g = 1$, on en déduit l'état de $H_0 + V$ au premier ordre en V :

$$\boxed{|\psi_n\rangle = |\psi_n^{(0)}\rangle + \sum_{m\neq n} \frac{\langle\psi_m^{(0)}|V|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}\,|\psi_m^{(0)}\rangle\ +\ \text{corrections d'ordre supérieur}} \tag{23.101}$$

Cette expression fournit une indication sur la validité de la méthode. En pratique, celle-ci est utile si les corrections restent petites, indiquant que les corrections successives décroissent rapidement. L'expression (23.101) montre que c'est le cas si les éléments de matrice de V restent petits devant les différences d'énergie[26], d'où les inégalités typiques à satisfaire pour que la méthode fonctionne convenablement :

$$\boxed{|\langle\psi_m^{(0)}|V|\psi_n^{(0)}\rangle| \ll |E_n^{(0)} - E_m^{(0)}|} \tag{23.102}$$

[26]L'expression de la correction d'énergie au second ordre, (23.106), confirme cette affirmation : le rapport entre un élément de matrice de perturbation et une différence d'énergie à l'ordre zéro est la *brique* élémentaire de tout développement en série de puissances.

À l'inverse, on voit qu'une perturbation ne produit d'effet notable que pour les couples d'états tels que :

$$|\langle\psi_m^{(0)}|V|\psi_n^{(0)}\rangle| \lesssim |E_n^{(0)} - E_m^{(0)}| \ ; \tag{23.103}$$

cette condition signifie que seuls les états séparés d'une énergie comparable (ou inférieure) à la perturbation sont sensiblement affectés par cette dernière. Elle constitue un critère permettant de tronquer *a priori* l'espace des états de façon raisonnable quand il s'agit de simplifier un problème. En outre, on voit ici pourquoi il est nécessaire d'envisager un traitement spécial pour le cas où le niveau non-perturbé présente une dégénérescence : alors, les dénominateurs d'énergie peuvent s'annuler et il convient de tout reprendre.

Une fois obtenues toutes les corrections relatives au premier ordre en g, on peut passer à l'ordre suivant. Il faut maintenant résoudre (23.83) où les inconnues sont ε_{n2} et $|\phi_{n2}\rangle$. ε_{n2} s'obtient immédiatement en multipliant à gauche par P_n. À nouveau le premier membre est nul ; après simplification du second membre, il reste :

$$0 = -P_n V|\phi_{n1}\rangle + \varepsilon_{n2}|\psi_n^{(0)}\rangle \tag{23.104}$$

d'où, après multiplication à gauche par $\langle\psi_n^{(0)}|$:

$$\boxed{\varepsilon_{n2} = \langle\psi_n^{(0)}|V|\phi_{n1}\rangle} \tag{23.105}$$

La correction du second ordre à l'énergie s'exprime à l'aide de la fonction perturbée au premier ordre, comme anticipé. En reportant l'expression (23.100), on en déduit :

$$\varepsilon_{n2} = \sum_{m\neq n} \frac{\langle\psi_n^{(0)}|V|\psi_m^{(0)}\rangle\,\langle\psi_m^{(0)}|V|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} = \sum_{m\neq n} \frac{|\langle\psi_n^{(0)}|V|\psi_m^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}} \ . \tag{23.106}$$

Il est utile de remarquer que cette correction du second ordre s'écrit de façon compacte à l'aide du projecteur Q_n :

$$\varepsilon_{n2} = \langle\psi_n^{(0)}|V[Q_n(E_n^{(0)} - H_0)Q_n]^{-1}V|\psi_n^{(0)}\rangle \ . \tag{23.107}$$

Cette correction peut être positive ou négative, puisque le dénominateur dans (23.106) n'est pas d'un signe donné. Toutefois, s'il s'agit de la correction pour l'état fondamental ($n = 1$, conventionnellement), tous les dénominateurs sont négatifs. On en déduit que pour l'état fondamental la correction du second ordre à l'énergie est toujours *négative* :

$$\boxed{\varepsilon_{n=1\,2} = -\sum_{m\neq n} \frac{|\langle\psi_1^{(0)}|V|\psi_m^{(0)}\rangle|^2}{E_m^{(0)} - E_1^{(0)}} < 0} \tag{23.108}$$

Il en résulte que, quand la correction au premier ordre est nulle, le déplacement d'énergie du fondamental est toujours vers le bas, comme si tous les niveaux excités *pesaient* sur le fondamental pour l'abaisser – bien sûr la même observation vaut pour le premier état excité si la perturbation isole l'état fondamental. Il en ira ainsi par exemple pour les

corrections radiatives[27] de l'état fondamental d'un atome couplé au rayonnement (dans une description où le champ est lui aussi quantifié). En ce qui concerne la correction du second ordre pour $|\psi_n\rangle$, $|\phi_{n2}\rangle$, elle est facile à écrire formellement ; repartant de (23.83), et multipliant à gauche par Q_n, on obtient :

$$|\phi_{n2}\rangle = [Q_n(H_0 - E_n^{(0)})Q_n]^{-1}Q_nVQ_n[Q_n(H_0 - E_n^{(0)})Q_n]^{-1}\, V|\psi_n^{(0)}\rangle \ ; \qquad (23.109)$$

la comparaison avec (23.98) indique clairement le mode de construction des différentes corrections $|\phi_{n\,k}\rangle$. Cette *mécanique* sera plus retrouvée plus tard (sous-section 23.2.3).

En résumé, le développement de l'énergie perturbée pour $H = H_0 + V$, quand il existe, commence comme suit :

$$E_n = E_n^{(0)} + \langle\psi_n^{(0)}|P_nVP_n|\psi_n^{(0)}\rangle + \langle\psi_n^{(0)}|V[Q_n(E_n^{(0)} - H_0)Q_n]^{-1}V|\psi_n^{(0)}\rangle + \ldots \ , \qquad (23.110)$$

ou, sous forme explicite :

$$E_n = E_n^{(0)} + \langle\psi_n^{(0)}|V|\psi_n^{(0)}\rangle + \sum_{m\neq n} \frac{|\langle\psi_n^{(0)}|V|\psi_m^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}} + \ldots \ . \qquad (23.111)$$

♦ *Remarque*

On a supposé que le niveau d'intérêt (identifié par n) est non-dégénéré, mais rien n'a été dit à ce sujet pour les autres niveaux, qui peuvent bien sûr présenter de la dégénérescence sans pour autant altérer la validité des résultats ci-dessus. Lorsqu'une telle dégénérescence existe, soit $\{|\psi_{m\alpha}^{(0)}\rangle\}_\alpha$ l'ensemble des états correspondants, tous associés à la même valeur de l'énergie, $E_m^{(0)}$, l'indice α allant de 1 à g_m, ordre de dégénérescence du niveau d'énergie $E_m^{(0)}$. Dans les formules ci-dessus, l'indice m est une étiquette pour tous les autres états en-dehors de n ; en présence de dégénérescence, les mêmes dénominateurs d'énergie vont se retrouver ici et là dans la sommation et on peut les regrouper, ce qui revient à remplacer la somme sur m par une somme sur $(p,\,\alpha)$ où p repère les énergies toutes différentes les unes des autres et où α distingue les états dégénérés entre eux. Ainsi, la correction du second ordre obtenue ci-dessus s'écrit tout autant :

$$\varepsilon_{n2} = \sum_{p\neq n} \frac{\sum_{\alpha=1}^{\mathrm{g}_p} |\langle\psi_{p\alpha}^{(0)}|V|\psi_n^{(0)}\rangle|^2}{E_n^{(0)} - E_p^{(0)}} \equiv \sum_{p\neq n}\sum_{\alpha=1}^{\mathrm{g}_p} \frac{|\langle\psi_{p\alpha}^{(0)}|V|\psi_n^{(0)}\rangle|^2}{E_p^{(0)} - E_m^{(0)}} \ . \qquad (23.112)$$

♦

[27]Le déplacement de l'énergie d'un niveau atomique sous l'effet de l'émission-absorption de photons (virtuels) s'appelle le déplacement de Lamb (*Lamb shift*) (voir p. 898) ; le *Lamb shift* de l'état fondamental est toujours négatif.

23.2.2 Cas d'un niveau dégénéré

Il est bien clair que l'hypothèse de non-dégénérescence joue un rôle très important dans les formules obtenues ci-dessus, où de multiples dénominateurs d'énergie apparaissent. Le cas d'un niveau dégénéré à l'ordre zéro doit donc faire l'objet d'un traitement spécial. En fait, la condition exprimée en (23.103) montre que les résultats doivent déjà être reconsidérés dès qu'il existe des énergies d'ordre zéro proches les unes des autres, cette proximité devant être appréciée à l'aune de l'ordre de grandeur des éléments de matrice de la perturbation, comme le montre (23.103). Cette situation se rencontre lorsqu'un croisement de niveaux se prépare : atome en champ magnétique assez fort (voir problème 23.4.11), faillite de l'approximation de Born - Oppenheimer (chapitre 28), proximité des lois de dispersion pour un électron sur réseau (chapitre 30, sous-section 30.4.1), etc. Dans la suite, le traitement d'une stricte dégénérescence est exposé, définissant une méthode qui doit en fait également être mise en œuvre pour les situations de *quasi*-dégénérescence.

Quans le niveau d'intérêt est dégénéré, il faut utiliser une notation précise, afin de pouvoir distinguer entre eux les états propres associés à une même valeur de l'énergie ; pour la suite, on adopte la notation suivante :

$$H_0|\psi_{n\alpha}^{(0)}\rangle = E_n^{(0)}\,|\psi_{n\alpha}^{(0)}\rangle \qquad (\alpha = 1,\, 2,\, \ldots,\, \mathrm{g}_n)\ . \tag{23.113}$$

L'ordre de dégénérescence, g_n, est supposé fini quelque soit n ; les $\{|\psi_{n\alpha}^{(0)}\rangle\}_\alpha$ engendrent le sous-espace propre $\mathcal{E}_n$ associé à $E_n^{(0)}$, de dimension g_n. Cela étant, n'importe quelle combinaison linéaire de ces g_n vecteurs satisfait l'équation aux vecteurs propres (23.113) :

$$\forall c_\alpha \in \mathbb{C}\ :\ H_0\sum_{\alpha=1}^{\mathrm{g}_n} c_\alpha|\psi_{n\alpha}^{(0)}\rangle = \sum_{\alpha=1}^{\mathrm{g}_n} c_\alpha H_0|\psi_{n\alpha}^{(0)}\rangle = \sum_{\alpha=1}^{\mathrm{g}_n} c_\alpha E_n^{(0)}|\psi_{n\alpha}^{(0)}\rangle = E_n^{(0)}\sum_{\alpha=1}^{\mathrm{g}_n} c_\alpha|\psi_{n\alpha}^{(0)}\rangle\ . \tag{23.114}$$

Bien évidemment, si l'on part de g fini et que l'on prend la limite $g \to 0$, il y a g_n vecteurs perturbés qui vont se réduire, par continuité, à *certaines* combinaisons linéaires bien déterminées des $|\psi_{n\alpha}^{(0)}\rangle$, mais on ne connaît pas ces combinaisons linéaires particulières, notées $|\phi_{n0\alpha}\rangle$. En présence de dégénérescence, on ne connaît pas d'avance les états propres à l'ordre zéro – au sens : états vers lequels tendent les états propres de H quand gV tend vers zéro. Ces états sont *a priori* différents des $|\psi_{n\alpha}^{(0)}\rangle$ fixés d'emblée pour une raison ou une autre : on sait seulement que tous ces vecteurs appartiennent au même sous-espace vectoriel $\mathcal{E}_n$, par définition[28].

Pour illustrer ce point, soit un espace de dimension 2, engendré par deux vecteurs $|\psi_{n1}^{(0)}\rangle$, $|\psi_{n2}^{(0)}\rangle$, tous deux propres d'un Hamiltonien H_0 avec la même valeur propre $E_n^{(0)}$. Sur cette base, la matrice de H_0 est :

$$H_0 = \begin{bmatrix} E_n^{(0)} & 0 \\ 0 & E_n^{(0)} \end{bmatrix}\ . \tag{23.115}$$

[28] En l'absence de dégénérescence, une telle question ne se pose pas : il n'y a qu'un seul vecteur perturbé $|\psi_n\rangle$ et il tend forcément vers $|\psi_n^{(0)}\rangle$.

Soit maintenant une perturbation gV dont la matrice, sur la même base, est :

$$V = \begin{bmatrix} 0 & gv \\ gv & 0 \end{bmatrix} . \tag{23.116}$$

Toute combinaison linéaire quelconque des deux vecteurs $|\psi_{n\alpha=1,2}^{(0)}\rangle$ est encore vecteur propre de H_0 avec toujours la même valeur propre $E_n^{(0)}$. D'un autre côté, les valeurs propres (ici exactes) de $H = H_0 + gV$ sont $E_{n\pm} = E_n^{(0)} \pm gv$, les vecteurs propres normalisés correspondants étant :

$$|\phi_{n\pm}\rangle = \frac{1}{\sqrt{2}}(\pm|\psi_{n1}^{(0)}\rangle + |\psi_{n2}^{(0)}\rangle) . \tag{23.117}$$

Maintenant, quand $g \to 0$, $E_{n\pm} \to E_n^{(0)}$, mais les deux vecteurs $|\phi_{n\pm}\rangle$ (qui ici ne dépendent[29] même pas de g) restent ce qu'ils sont ; les combinaisons linéaires particulières, limites (et pour cause) des états propres de H quand la perturbation tend vers zéro, n'ont aucune raison de coïncider avec les $|\psi_{n\alpha=1,2}^{(0)}\rangle$ préalablement définis – ils sont pourtant bien, à la limite, des états propres de H_0, au même titre d'ailleurs que *toute* autre combinaison linéaire.

Revenant au cas général, on voit qu'il existe donc exactement g_n vecteurs perturbés, $|\psi_{n\alpha}\rangle$, qui vont tendre vers des combinaisons linéaires $|\phi_{n0\alpha}\rangle$ à déterminer ; pour chacun de ces $|\psi_{n\alpha}\rangle$, on suppose toujours qu'il est possible d'écrire un développement en puissances entières de g. À chacun de ces vecteurs correspond une énergie perturbée E_n ; les E_n seront ou ne seront pas toutes distinctes. Si elles sont toutes différentes, on dit que la perturbation a complètement levé toute la dégénérescence ; dans le cas contraire, la levée de dégénérescence est seulement partielle. Quoi qu'il en soit, on a maintenant, repérant toujours par n le niveau auquel on s'intéresse particulièrement :

$$E_{n\alpha} = E_n^{(0)} + g\varepsilon_{n1\alpha} + g^2\varepsilon_{n2\alpha} + \ldots , \quad |\psi_{n\alpha}\rangle = |\phi_{n0\alpha}\rangle + g|\phi_{n1\alpha}\rangle + g^2|\phi_{n2\alpha}\rangle + \ldots \tag{23.118}$$

Comme précédemment, les corrections $|\phi_{nk\alpha}\rangle$, $k \geq 1$ peuvent être choisies orthogonales à tous les $|\psi_{n\alpha}^{(0}\rangle$ (donc également aux $|\phi_{n0\alpha}\rangle$), c'est-à-dire appartenant au complément orthogonal de $\mathcal{E}_n$. En reportant ces développements dans l'équation aux vecteurs propres et en identifiant les termes de même puissance en g, on trouve successivement :

$$H_0|\phi_{n0\alpha}\rangle = E_n^{(0)}|\phi_{n0\alpha}\rangle , \tag{23.119}$$

$$H_0|\phi_{n1\alpha}\rangle + V|\phi_{n0\alpha}\rangle = E_n^{(0)}|\phi_{n1\alpha}\rangle + \varepsilon_{n1\alpha}|\phi_{n0\alpha}\rangle , \tag{23.120}$$

$$H_0|\phi_{n2\alpha}\rangle + V|\phi_{n1\alpha}\rangle = E_n^{(0)}|\phi_{n2\alpha}\rangle + \varepsilon_{n1\alpha}|\phi_{n1\alpha}\rangle + \varepsilon_{n2\alpha}|\phi_{n0\alpha}\rangle , \tag{23.121}$$

et ainsi de suite. À nouveau, la première équation n'apprend rien. La seconde, issue du terme en g, s'écrit :

$$(H_0 - E_n^{(0)})|\phi_{n1\alpha}\rangle = (\varepsilon_{n1\alpha} - V)|\phi_{n0\alpha}\rangle . \tag{23.122}$$

[29]Ceci vient du fait que l'exemple est trop simple, et que les solutions exactes correspondent à une *symétrie* donnée, brisée dès que $g \neq 0$, aussi petit soit-il.

Introduisons maintenant le projecteur P_n sur le sous-espace propre associé à $E_n^{(0)}$:

$$P_n \stackrel{\text{déf}}{=} \sum_{\alpha=1}^{g_n} |\psi_{n\alpha}^{(0)}\rangle\langle\psi_{n\alpha}^{(0)}| = \sum_{\alpha=1}^{g_n} |\phi_{n0\alpha}\rangle\langle\phi_{n0\alpha}| \ . \tag{23.123}$$

la deuxième égalité venant du fait que les $|\psi_{n\alpha}^{(0)}\rangle$ et les $|\phi_{n0\alpha}\rangle$ engendrent le *même* sous-espace $\mathcal{E}_n$; on a bien sûr $P_n|\psi_{n\alpha}^{(0)}\rangle = |\psi_{n\alpha}^{(0)}\rangle$, $P_n|\phi_{n0\alpha}\rangle = |\phi_{n0\alpha}\rangle$. Le projecteur dans l'espace orthogonal est :

$$Q_n = \mathbf{1} - P_n = \sum_{m\neq n} \sum_{\beta=1}^{g_m} |\psi_{m\beta}^{(0)}\rangle\langle\psi_{m\beta}^{(0)}| \equiv \sum_{m\neq n} P_m \ . \tag{23.124}$$

Par définition de Q_n, $Q_n|\psi_{n\alpha}^{(0)}\rangle = Q_n|\phi_{n0\alpha}\rangle = 0$; noter que P_n n'est autre que l'identité dans le sous-espace dégénéré $\mathcal{E}_n$. En multipliant (23.122) à gauche par P_n :

$$P_n(H_0 - E_n^{(0)})|\phi_{n1\alpha}\rangle = P_n(\varepsilon_{n1\alpha} - V)|\phi_{n0\alpha}\rangle \ . \tag{23.125}$$

Le premier membre est nul, puisque $P_nH_0 - E_n^{(0)}P_n = 0$ (et d'ailleurs les $|\phi_{n1\alpha}\rangle$ sont dans le complément orthogonal), d'où $P_n(\varepsilon_{n1\alpha} - V)|\phi_{n0\alpha}\rangle = 0$. Injectant $\mathbf{1} = P_n + Q_n$ juste devant $|\psi_{n0\alpha}^{(0)}\rangle$ et avec $Q_n|\phi_{n0\alpha}\rangle = 0$, il vient $P_n(V - \varepsilon_{n1\alpha})P_n|\phi_{n0\alpha}\rangle = 0$, d'où :

$$\boxed{(P_nVP_n)|\phi_{n0\alpha}\rangle = \varepsilon_{n1\alpha}|\phi_{n0\alpha}\rangle} \tag{23.126}$$

Cette équation établit de fait un résultat important : elle montre que $|\phi_{n0\alpha}\rangle$ est vecteur propre de P_nVP_n avec la valeur propre $\varepsilon_{n1\alpha}$; les corrections au premier ordre $\varepsilon_{n1\alpha}$ ne sont donc rien d'autre que les valeurs propres de l'opérateur P_nVP_n, et les bons états à l'ordre zéro[30] – dont on rappelle qu'ils étaient eux aussi inconnus – sont les vecteurs propres associés. L'équation (23.126) définit *ipso facto* la méthode de perturbation au premier ordre pour un niveau dégénéré : on écrit la matrice de V dans le seul sous-espace $\mathcal{E}_n$, sans se soucier de ce qu'elle est ailleurs, et on la diagonalise. Notons que le résultat (23.93) n'est qu'un cas particulier quand $g_n = 1$ (pas de dégénérescence).

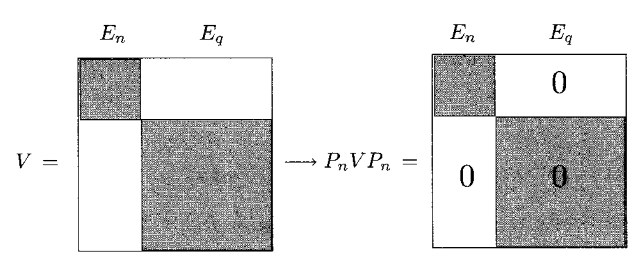


Figure 23.1: Illustration de la projection de l'opérateur de perturbation dans le sous-espace dégénéré $\mathcal{E}_n$.

[30]au sens : limite des états perturbés lorsque la perturbation tend vers zéo.

L'opérateur $P_n V P_n$ est la *projection* de l'opérateur de perturbation V dans le sous-espace propre $\mathcal{E}_n$, puisque les éléments de matrice de $P_n V P_n$ sont tels que :

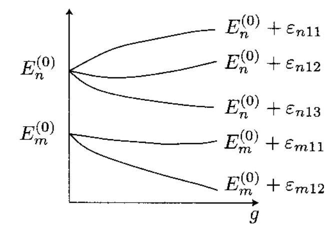
$$\langle\psi_{n\alpha}^{(0)}|P_n V P_n|\psi_{n\alpha'}^{(0)}\rangle = \langle\psi_{n\alpha}^{(0)}|V|\psi_{n\alpha'}^{(0)}\rangle,\ \forall\,\alpha,\ \alpha'\ ,\ \langle\psi_{n\alpha}^{(0)}|V|\psi_{m\beta}^{(0)}\rangle = 0\,, \forall n \neq m,\ \forall\alpha,\ \beta\ . \tag{23.127}$$

Autrement dit, ils coïncident avec ceux de V dans le sous-espace $\mathcal{E}_n$ et sont tous nuls partout ailleurs ; la figure 23.1 est une illustration schématique de la projection, en terme de la matrice de V.

En considérant successivement tous les sous-espaces dégénérés, on voit que la résolution du problème donnant les états à l'ordre zéro et les valeurs propres à l'ordre 1 implique non pas l'opérateur V présent au départ mais la somme (directe) de ses projections dans chaque sous-espace $\mathcal{E}_n$; algébriquement parlant : on remplace V par un opérateur approché, V_{ap}, tel que :

$$V_{\text{ap}} \stackrel{\text{déf}}{=} \sum_n P_n V P_n \neq V = \sum_n P_n V P_n + \sum_{n\neq m} (P_n V P_m + P_m V P_n)\ , \tag{23.128}$$

où tous les blocs non-diagonaux connectant deux sous-espaces dégénérés distincts dans V ont été annulés identiquement dans V_{ap}.



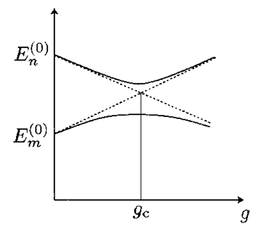


Figure 23.2: À gauche : illustration de la levée de dégénérescence au premier ordre par une perturbation. À droite : croisement de niveaux évité, situation dans laquelle il faut employer la technique de perturbation dégénérée dans le *voisinage* de g_c, et non pas seulement en $g = g_c$.

En résumé, la méthode de résolution *minimale* pour un niveau dégénéré consiste à diagonaliser la matrice de la perturbation V, considérée seulement dans le sous-espace dégénéré et en ignorant tout le reste[31] ; les valeurs propres ainsi obtenues sont les premières corrections à l'énergie. L'opérateur V étant hermitique, sa projection l'est tout autant et toutes les corrections $\varepsilon_{n1\alpha}$ sont réelles ; les énergies correspondantes $E_n^{(0)} + g\,\varepsilon_{n1\alpha}$ sont les approximations des E_n. La perturbation a donc, en général, la vertu de scinder en plusieurs niveaux distincts les g_n états dégénérés en l'absence de V (voir fig. 23.2) ; la levée de dégénérescence est *complète* si l'on obtient exactement g_n énergies distinctes.

[31] On imagine le rôle de premier plan joué en pareille circonstance par le théorème de Wigner - Eckart, puisque celui-ci énonce un résultat concernant un sous-espace *donné* $|\alpha j m\rangle$.

Pour aller plus loin dans les corrections, il convient de faire le tri : au premier ordre pour l'énergie, certains états sont encore dégénérés, d'autres ne le sont plus, quand la levée est partielle. Les derniers peuvent être alors traités par le schéma simple relatif à un niveau non-dégénéré ; les autres au contraire relèvent du traitement qui vient d'être exposé. Il est clair que la discussion générale des corrections d'ordre supérieur est sans intérêt mais que celles-ci sont calculables au coup par coup, parfois laborieusement.

Comme dit plus haut, la distinction entre niveaux dégénérés et niveaux non-dégénérés n'est pas toujours franche, et doit en fait être appréciée en comparant les éléments de matrice de la perturbation avec les différences d'énergie à l'ordre zéro. En cas de quasi-dégénérescence, au sens de (23.103), la sûreté exige d'appliquer la technique élaborée pour le cas dégénéré. En outre, il arrive que les niveaux dépendent continûment d'un paramètre (le module d'un champ extérieur, par exemple) et que, ce paramètre variant, ils en viennent à se croiser (croisement de niveaux). En pareil cas, au *voisinage* du point de croisement (et pas seulement au point de croisement !) il convient de traiter le problème dans un cadre unique – indépendant de la valeur du paramètre – et c'est à nouveau le cadre de la méthode de la situation dégénérée qu'il faut choisir : diagonaliser la matrice de V réduite au sous-espace des états pouvant se croiser (voir problème 23.4.11). Selon les cas (et notamment la symétrie[32] de V), le point de croisement peut diparaître – on parle alors de *croisement évité* (fig. 23.2 à droite) – ou au contraire persister.

23.2.3 Développement systématique à l'aide de la résolvante

La résolvante $G(z)$ est définie comme l'opérateur[33] :

$$\boxed{G(z) \stackrel{\text{déf}}{=} \frac{1}{z\mathbf{1} - H}} \tag{23.129}$$

Dans toute la suite, pour simplifier l'écriture, on omettra $\mathbf{1}$ en facteur de tous les scalaires. z est un nombre complexe quelconque. $G(z)$ est relié simplement à la transformée de Laplace de l'opérateur d'évolution avancé $U_+(t)$:

$$U_+(t) \stackrel{\text{déf}}{=} \begin{cases} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Ht} & \text{si } t > 0 \\ 0 & \text{si } t < 0 \end{cases} , \qquad U_+(t=0) = \mathbf{1} \ . \tag{23.130}$$

En effet, soit $K(z)$ le transformé de Laplace de l'opérateur $U_+(t)$ défini par (23.130) :

$$K(z) \equiv \mathcal{L}(U_+) \stackrel{\text{déf}}{=} \int_0^{+\infty} \mathrm{e}^{-zt}\, \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}Ht}\, \mathrm{d}t \ ; \tag{23.131}$$

l'intégrale peut s'exprimer comme avec des scalaires et on obtient :

$$K(z) = \frac{1}{z\mathbf{1} - (H/\mathrm{i}\hbar)} \equiv \mathrm{i}\hbar\, G(\mathrm{i}\hbar z) \ , \tag{23.132}$$

[32] Si l'élément de matrice de V est nul par symétrie, le point de croisement demeure.

[33] Tout ce qui suit n'a évidemment de sens que pour un Hamiltonien H indépendant du temps.

soit, à l'envers :

$$G(z) = (\mathrm{i}\hbar)^{-1}\, K(-\mathrm{i}\hbar^{-1}z) \ . \tag{23.133}$$

Explicitement, on a la relation intégrale inverse :

$$\boxed{U_+(t) = \frac{1}{2\mathrm{i}\pi}\int_{\mathrm{C}_+} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\,zt}\, G(z)\,\mathrm{d}z} \tag{23.134}$$

où C_+ est un contour situé juste au-dessus de l'axe réel, parcouru dans le sens des parties réelles décroissantes, de $+\infty$ à $-\infty$.

L'intérêt de la résolvante est sa forme algébrique simple qui se prête remarquablement aux développements comme on va le voir. En outre, les propriétés analytiques de $G(z)$ traduisent fidèlement la nature du spectre de H et des arguments généraux utilisant l'arsenal de la Théorie des fonctions analytiques permettent souvent d'obtenir des résultats qualitatifs de première importance sans effectuer de lourds calculs. En effet, supposant que le spectre est composé d'un ensemble discret $\{E_n\}_n$ et d'une portion continue allant de E_b à l'infini (par exemple)[34], $G(z)$ prend la forme explicite :

$$G(z) = \sum_n \frac{P_n}{z - E_n} + \int_{E_\mathrm{b}}^{+\infty} \frac{P(E)}{z - E}\,\mathrm{d}E \tag{23.135}$$

où $P_n = \sum_\alpha |E_{n\alpha}\rangle\langle E_{n\alpha}|$ et $P(E) = \int |E,\,\alpha\rangle\langle E,\,\alpha|\,\mathrm{d}\alpha$ sont les projecteurs sur les états propres de H – la trace de P_n est la dimension g_n du sous-espace dégénéré associé à l'énergie E_n. Ainsi, $G(z)$ a un pôle simple pour chaque valeur discrète du spectre, et une coupure sur l'intervalle continu $[E_\mathrm{b},\,+\infty[$: les propriétés analytiques de $G(z)$ sont donc le reflet fidèle du spectre du Hamiltonien. Notons que l'on peut facilement inverser la relation (23.135) grâce au théorème des résidus ; par exemple :

$$\boxed{P_n = \frac{1}{2\mathrm{i}\pi}\oint_{\mathrm{C}_n} G(z)\,\mathrm{d}z} \tag{23.136}$$

où C_n est un contour fermé quelconque, parcouru une fois dans le sens positif, contenant exclusivement le pôle E_n. Il est clair que connaître les modes propres de H ou ceux de $G(z)$ est équivalent, à des difficultés techniques près[35].

Dans le cas où $H = H_0 + gV$, on peut définir deux résolvantes $G_0 = (z - H_0)^{-1}$ et $G(z) = (z - H)^{-1}$. Ceci étant posé, il est alors facile d'établir une équation fondamentale, dite *équation de Dyson*, reliant G et G_0. Écrivons :

$$z\mathbf{1} - H_0 = z\mathbf{1} - H_0 - gV + gV = z\mathbf{1} - H + gV \ , \tag{23.137}$$

soit $G_0^{-1} = G^{-1} + gV$, égalité qui, multipliée à gauche par G_0, donne $\mathbf{1} = G_0\,G^{-1} + g\,G_0 V$; enfin, multipliant à droite par G, on obtient l'équation de Dyson :

$$\boxed{G(z) = G_0(z) + g\,G_0(z)\,V\,G(z)} \tag{23.138}$$

[34] Le plus souvent, on a de surcroît : $\max_n\{E_n\} < E_\mathrm{b}$.

[35] $G(z)$ est d'un emploi très commode ; il reste que le retour à H, *via* $U(t)$ et ses fréquences d'oscillation, nécessite une transformation de Laplace inverse qui n'est pas toujours aisée.

qui est une équation exacte mais implicite puisque l'inconnue G figure dans les deux membres.

$G(z)$ $G_0(z)$ $G(z)$ $G_0(z)$ $G_0(z)$ $G(z)$

━━━ = ─── V ━━━ = ─── V ─── V ━━━ = ⋯

Figure 23.3: Représentation imagée de l'équation de Dyson (23.138) et de l'égalité (23.141).

L'équation (23.138) s'itère formellement comme suit :

$$G = G_0 + gG_0V(G_0 + gG_0VG) \; ; \tag{23.139}$$

répétant la substitution à l'infini, on en déduit le développement exact :

$$G = G_0 + g\, G_0VG_0 + g^2\, G_0VG_0VG_0 + \ldots \; . \tag{23.140}$$

soit :

$$\boxed{G = G_0 + G_0 \sum_{k=1}^{+\infty} g^k \,(VG_0)^k} \tag{23.141}$$

Ce développement, tronqué à un certain ordre, est équivalent à une théorie de perturbation au même ordre ([15], p. 611). De façon plus explicite, la méthode consiste à partir de (23.136) – où C_n est un contour entourant maintenant la valeur propre non perturbée $E_n^{(0)}$ *et*[36] celles qui tendent vers elle lorque $g \to 0$ – et à essayer de calculer les différents termes :

$$P_n = \frac{1}{2\mathrm{i}\pi} \oint_{\mathrm{C}_n} \left[G_0(z) + gG_0(z)VG_0(z) + g^2G_0(z)VG_0(z)VG_0(z) + \ldots\right] \mathrm{d}z \; . \tag{23.142}$$

On en déduit alors l'expression des projecteurs, P_n par exemple, sous la forme d'un développement en puissances de g :

$$P_n = P_n^{(0)} + \sum_{k=1}^{+\infty} \frac{g^k}{2\mathrm{i}\pi} \oint_{\mathrm{C}_n} G_0(z)\left[VG_0(z)\right]^k \mathrm{d}z \; . \tag{23.143}$$

À l'intérieur de C_n, chaque terme de la série a pour seule singularité un pôle d'ordre $k+1$ en $E_n^{(0)}$. Pour un niveau non-dégénéré, l'énergie s'obtient en partant de :

$$\boxed{E_n = \mathrm{Tr}(HP_n)} \tag{23.144}$$

puisque $HP_n = E_nP_n$ et que $\mathrm{Tr}P_n = 1$. Formellement :

$$E_n = \mathrm{Tr}\Big\{(H_0 + gV)\Big[P_n^{(0)} + \sum_{k=1}^{+\infty} \frac{g^k}{2\mathrm{i}\pi} \oint_{\mathrm{C}_n} \mathrm{d}z\, G_0(z)[VG_0(z)]^k\Big]\Big\} \tag{23.145}$$

[36] faute de quoi chaque terme de la série dans (23.142) aurait un résidu nul.

soit : $E_n = \tilde{E}_n^{(0)} + \sum_{k=1}^{+\infty} g^k \,\varepsilon_{nk}$ avec :

$$\tilde{E}_n^{(0)} = \mathrm{Tr}\{(H_0 + gV)P_n^{(0)}\} \ , \tag{23.146}$$

$$\varepsilon_{nk} = \frac{1}{2\mathrm{i}\pi} \oint_{\mathrm{C}_n} \mathrm{d}z \, \mathrm{Tr}\{(H_0 + gV)G_0(z)[VG_0(z)]^k\} \ . \tag{23.147}$$

$\tilde{E}_n^{(0)}$ est l'énergie *renormalisée* incorporant les éventuels éléments diagonaux de V (que l'on peut d'ailleurs inclure dans H_0 en le redéfinissant).

Pour terminer, montrons sur un exemple simple un autre intérêt de la résolvante. Soit une perturbation V définie comme[37] :

$$V = v|\psi\rangle\langle\psi| \quad (v \in \mathbb{R}) \tag{23.148}$$

où $|\psi\rangle$ est un vecteur donné de l'espace des états. Soit à calculer l'élément diagonal $\langle\phi|G(z)|\phi\rangle$ où $|\phi\rangle$ est *a priori* quelconque ; l'équation itérée (23.141) permet d'écrire :

$$\langle\phi|G|\phi\rangle = \langle\phi|G_0|\phi\rangle + v\langle\phi|G_0|\psi\rangle\langle\psi|G_0|\phi\rangle + v^2\langle\phi|G_0|\psi\rangle\langle\psi|G_0|\psi\rangle\langle\psi|G_0|\phi\rangle + \ldots \tag{23.149}$$

Posons $f_0(z) \stackrel{\text{déf}}{=} \langle\phi|G_0(z)|\phi\rangle$, $h(z) \stackrel{\text{déf}}{=} \langle\phi|G_0(z)|\psi\rangle\langle\psi|G_0|\phi\rangle$, $k(z) \stackrel{\text{déf}}{=} \langle\psi|G_0(z)|\psi\rangle$; avec ces définitions :

$$\langle\phi|G(z)|\phi\rangle = f_0(z) + v\,h(z)h^*(z) + v^2 h(z)k(z)h^*(z) + v^3 h(z)[k(z)]^2 h^*(z) + \ldots \tag{23.150}$$

où h^* doit être compris comme la fonction complexe conjuguée de h lorsque z est réel ; la série géométrique se resomme immédiatement[38] :

$$\langle\phi|G(z)|\phi\rangle = f_0(z) + vh(z)\Big[1 + vk(z) + v^2[k(z)]^2 + \ldots\Big]h^*(z) = f_0(z) + v\,\frac{h(z)h^*(z)}{1 - vk(z)} \tag{23.151}$$

et donne une expression compacte et exacte de l'élément de matrice de G, resommée à tous les ordres de la théorie de perturbation. Le même problème, posé en théorie de perturbation standard aurait fourni des expressions tronquées à un ordre donné et n'aurait sûrement pas fourni le dénominateur présent dans (23.151), qui donne à $\langle\phi|G(z)|\phi\rangle$ ses propriétés analytiques : la puissance de la résolvante peut se mesurer à l'aune de cette dernière remarque. Plus généralement, la résolvante se prête bien à l'imagination de méthodes approchées nettement plus séduisantes que la théorie des perturbations standard, l'objectif étant au coup par coup de parvenir à fabriquer une résolvante (approchée) G_{ap} ayant la forme algébrique d'une résolvante[39], soit $G_{\text{ap}} = (z\mathbf{1} - H_{\text{ap}})^{-1}$, H_{ap} jouant le rôle d'un Hamiltonien effectif ; généralement parlant, une telle forme est obtenue par resommation à l'infini d'une ou plusieurs *classes* de termes de la série de perturbation.

[37] Cette perturbation peut par exemple représenter une impureté localisée en phase solide.

[38] On suppose bien sûr que $|vk(z)| < 1$.

[39] Un exemple de cette manière de faire est l'approximation dite *du pôle*, dont la portée est générale (elle conduit à la notion de *résonance*). Une application en est faite dans le chapitre 25, section 25.2, pour la description de l'émission spontanée d'un atome.

23.2.4 Approximation de l'équation aux valeurs propres sous sa forme intégrale

En représentation-q, l'équation aux valeurs et fonctions propres pour une particule de masse m dans le potentiel $V(\vec{r})$, supposé tendre vers zéro à l'infini, est :

$$-\frac{\hbar^2}{2m}\Delta\psi(\vec{r}) + V(\vec{r})\psi(\vec{r}) = E\psi(\vec{r}) \ ; \tag{23.152}$$

en posant $E = \frac{\hbar^2 k^2}{2m}$ et $U(\vec{r}) \stackrel{\text{déf}}{=} \frac{2m}{\hbar^2}V(\vec{r})$, elle prend la forme :

$$\boxed{\Delta\psi(\vec{r}) + k^2\psi(\vec{r}) = U(\vec{r})\psi(\vec{r})} \tag{23.153}$$

Considérons[40] cette équation comme une équation aux dérivées partielles *inhomogène* avec la source $U(\vec{r})\psi(\vec{r})$; la méthode de la fonction de Green permet alors d'écrire la solution de (23.153) sous la forme :

$$\psi(\vec{r}) = \psi_0(\vec{r}) + \int_{\mathbb{R}^3} G(\vec{r},\, \vec{r}\,')U(\vec{r}\,')\psi(\vec{r}\,')\,\mathrm{d}^3r' \ , \tag{23.154}$$

où $\psi_0(\vec{r})$ est solution de l'équation homogène $\Delta\psi_0(\vec{r}) + k^2\psi_0(\vec{r}) = 0$, et où la fonction de Green $G(\vec{r},\, \vec{r}\,')$ satisfait[41] :

$$\Delta_{\vec{r}}\, G(\vec{r},\, \vec{r}\,') + k^2 G(\vec{r},\, \vec{r}\,') = \delta(\vec{r} - \vec{r}\,') \ . \tag{23.155}$$

Le second membre étant invariant par translation, la fonction $G(\vec{r},\, \vec{r}\,')$ ne dépend que de la différence $\vec{r} - \vec{r}\,'$, et on la note désormais $G(\vec{r})$. Pour résoudre cette équation, on passe dans l'espace de Fourier en posant :

$$G(\vec{r}) = \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} \mathcal{G}(\vec{k}')\mathrm{e}^{\mathrm{i}\vec{k}'.\vec{r}}\,\mathrm{d}^3k' \ . \tag{23.156}$$

La transformée de Fourier de l'équation (23.155) est :

$$(-k'^2 + k^2)\mathcal{G}(\vec{k}') = 1 \ , \tag{23.157}$$

qui montre que $\mathcal{G}$ ne dépend que de $k' \equiv \|\vec{k}'\|$; exploitant cette propriété pour effectuer l'intégration sur les angles dans l'égalité (23.156), on obtient :

$$G(\vec{r}) = \frac{1}{2\pi^2 r}\int_0^{+\infty} \mathcal{G}(k')\sin k'r\, k'\mathrm{d}k' \equiv G(r) \ . \tag{23.158}$$

[40]On retrouvera plus tard strictement la même façon de procéder, à propos de la théorie de la diffusion par un potentiel – voir section 26.3.

[41]L'équation (23.155) montre que la dimension de la fonction de Green est ici L^{-1}. Dans $\mathbb{R}^D$, elle a pour dimension $[G] = \mathrm{L}^{2-D}$. La forme analytique précise de G dépend de la dimension spatiale – voir par exemple le cas $D = 1$, traité dans le problème 26.6.1.

Examinons d'abord le cas des états liés, c'est-à-dire correspondant à $E < 0$, soit k imaginaire pur, $k \equiv \mathrm{i}\kappa$ avec $\kappa = |k| \geq 0$. La solution de (23.157) est $\mathcal{G}(k') = -\frac{1}{k'^2+\kappa^2}$; le report dans (23.158) donne immédiatement par résidus[42] :

$$\boxed{G(r) = -\frac{1}{4\pi r}\,\mathrm{e}^{-\kappa r}} \tag{23.160}$$

Pour les états non-liés, k est réel (et on peut toujours supposer $k > 0$). Alors, l'équation (23.155) a pour solution particulière la partie principale de Cauchy $\mathcal{P}\frac{1}{k^2-k'^2}$, à laquelle on peut ajouter n'importe quelle fonction $\delta(k' \pm k)$ puisque $(k' \pm k)\delta(k' \pm k) = 0$. La solution générale s'écrit alors :

$$\mathcal{G}(k) = \mathcal{P}\frac{1}{k^2 - k'^2} + C_+\delta(k' - k) + C_-\delta(k' + k) \ , \tag{23.161}$$

où les constantes $C_\pm$ sont pour l'instant quelconques. Revenant à (23.158), on a :

$$G(\vec{r}) = \frac{1}{2\pi^2 r}\int_0^{+\infty}\left[\mathcal{P}\frac{1}{k^2 - k'^2} + C_+\delta(k' - k) + C_-\delta(k' + k)\right]\sin k'r\, k'\mathrm{d}k' \ . \tag{23.162}$$

La partie principale est égale à $\frac{1}{2\mathrm{i}}\int_{-\infty}^{+\infty}\mathcal{P}\frac{k'}{k^2-k'^2}\,\mathrm{e}^{\mathrm{i}k'r}\mathrm{d}k'$ et vaut $-\frac{\pi}{2}\cos kr$, d'où :

$$\boxed{G(r) = -\frac{1}{4\pi r}(\cos kr + C\sin kr)} \tag{23.163}$$

où C est une constante arbitraire[43].

Qu'il s'agisse des états liés ou non-liés, l'équation aux valeurs propres est donc équivalente à l'équation (23.154), munie de la fonction de Green appropriée, et est un exemple d'équation intégrale de Fredholm de seconde espèce. Si le problème reste entier puisque l'inconnue figure $\psi(\vec{r})$ dans les deux membres, la forme intégrale est parfois plus aisée et, en tout cas, suggère un schéma d'approximations successives en itérant la relation intégrale[44].

En ce qui concerne les états liés, la condition aux limites à l'infini est $\psi(\infty) = 0$; il faut donc en particulier $\psi_0(\infty) = 0$. La solution générale de l'équation homogène est une combinaison linéaire quelconque $A_+\mathrm{e}^{\kappa r} + A_-\mathrm{e}^{-\kappa r}$, et la seule fonction satisfaisant

[42] Reportant cette solution dans (23.155), on en déduit l'égalité utile à connaître :

$$\boxed{\Delta\frac{1}{r}\,\mathrm{e}^{-\kappa r} = \frac{\kappa^2}{r}\,\mathrm{e}^{-\kappa r} - 4\pi\delta(\vec{r}) \qquad (\kappa \geq 0)} \tag{23.159}$$

Quand $\kappa = 0$, c'est essentiellement l'équation de Poisson pour une charge ponctuelle.

[43] La deuxième constante C_- disparaît puisque k est conventionnellement défini comme positif, et que l'intégrale dans (23.158) est prise sur $\mathbb{R}_+$; quant à la constante C, elle sera déterminée dans chaque cas par l'application des bonnes conditions aux limites pour le problème considéré.

[44] On procédera exactement de la même façon dans le chapitre 24 à propos de l'opérateur d'évolution (voir section 24.2), et pour l'équation intégrale de la diffusion (chapitre 26, section 26.3).

la condition aux limites est celle pour laquelle $A_\pm = 0$: c'est la fonction identiquement nulle. Cela étant, l'équation de Fredholm (23.154) devient *homogène*[45] ; dans ce cas, il n'existe de solution non identiquement nulle que si le paramètre k satisfait une certaine équation [184], qui n'est rien d'autre qu'un avatar de l'équation donnant les valeurs propres E.

Pour les états non-liés (états de diffusion), le problème physique doit être spécifié sans ambiguïté en précisant la localisation de la source émettant les particules envoyées en direction du potentiel $V(\vec{r})$, ce qui revient à définir une certaine condition aux limites pour toute solution ; par exemple, si on choisit l'axe Oz le long de la direction du faisceau incident, et si la source est en $z = -\infty$, le terme inhomogène dans (23.154) est de la forme $\mathrm{e}^{\mathrm{i}kz}$ (k est toujours supposé positif) puisque c'est la solution en l'absence de potentiel et que, pour une telle localisation de la source et sans potentiel, la solution est *partout* égale à $\mathrm{e}^{\mathrm{i}kz}$, à un facteur près. Avec une telle géométrie, on peut montrer (voir chapitre 26, section 26.3) que la solution satisfaisant ces conditions aux limites dans $\mathbb{R}^3$ est :

$$\boxed{G(\vec{r}, \vec{r}\,') = -\frac{1}{4\pi\|\vec{r}-\vec{r}\,'\|}\,\mathrm{e}^{\mathrm{i}k\|\vec{r}-\vec{r}\,'\|}} \tag{23.164}$$

et l'équation aux fonctions propres (23.152) se retranscrit alors exactement :

$$\boxed{\psi(\vec{r}) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} - \frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{1}{\|\vec{r}-\vec{r}\,'\|}\,\mathrm{e}^{\mathrm{i}k\|\vec{r}-\vec{r}\,'\|}\,U(\vec{r}\,')\psi(\vec{r}\,')\,\mathrm{d}^3r'} \tag{23.165}$$

où $\vec{k}/k$ définit la direction du faisceau incident. Ceci constitue la forme intégrale de l'équation de Lippmann - Schwinger [79], qui joue un rôle de premier plan en théorie des collisions. Bien sûr, cette équation est *implicite* pour la fonction inconnue $\psi(\vec{r})$, mais l'idée qu'elle suggère ici est d'effectuer des approximations successives, en itérant cette relation intégrale et en s'arrêtant à un ordre donné par rapport au potentiel $V(\vec{r})$. Ainsi, à l'ordre le plus bas non-trivial, on remplace l'inconnue $\psi(\vec{r})$ par sa valeur à l'ordre zéro et on obtient :

$$\psi(\vec{r}) \simeq \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} - \frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{1}{\|\vec{r}-\vec{r}\,'\|}\,\mathrm{e}^{\mathrm{i}k\|\vec{r}-\vec{r}\,'\|}\,U(\vec{r}\,')\psi_0(\vec{r}\,')\,\mathrm{d}^3r'. \tag{23.166}$$

La solution à $V(\vec{r}) \equiv 0$ peut être prise sous la forme d'une onde plane ; normalisant celle-ci dans une boîte cubique de côté L *aussi grand que l'on veut*, la solution explicite à cet ordre en présence du potentiel est :

$$\psi(\vec{r}) \simeq \frac{1}{L^{3/2}}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} - \frac{1}{4\pi L^{3/2}}\int_{\mathbb{R}^3} \frac{1}{\|\vec{r}-\vec{r}\,'\|}\,\mathrm{e}^{\mathrm{i}k\|\vec{r}-\vec{r}\,'\|}\,U(\vec{r}\,')\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}\,'}\,\mathrm{d}^3r'. \tag{23.167}$$

Si a est la portée du potentiel $V(\vec{r})$, on voit tout de suite que pour $r \gg a$, l'intégrale est de la forme $S(\vec{r}/r)\,\frac{1}{r}\,\mathrm{e}^{\mathrm{i}kr}$ et décrit donc une onde sphérique *sortante* diffusée par le potentiel $V(\vec{r})$, le facteur $S(\vec{r}/r)$ contenant la dépendance angulaire, par rapport à

[45] Les deux cas états liés/états non-liés constituent un bel exemple physique du théorème dit *de l'alternative de Fredholm* [184].

la direction incidente, de l'onde diffusée et constituant l'image angulaire du potentiel ; quant au terme d'ordre zéro, il représente la diffusion vers l'avant, triviale et toujours présente. L'expression (23.167) constitue la *première approximation de Born* ; elle sera reprise et développée dans le chapitre 26, section 26.3. Clairement, il s'agit de la première étape d'un processus itératif d'approximation dont l'utilité en pratique est assujettie à la décroissance rapide des corrections successives.

23.3 Exemples d'application : effets Stark et Zeeman pour l'atome d'hydrogène

Il s'agit ici de traiter deux applications classiques de la théorie des perturbations, donnant une description élémentaire de deux effets importants : l'effet Stark[46] (éclatement des sous-niveaux atomiques par application d'un champ électrique) et l'effet Zeeman (phénomène analogue provoqué par un champ magnétique). Expérimentalement, ces effets se manifestent par l'apparition de raies très voisines là où, en l'absence de champ, on observe une raie unique.

23.3.1 Effet Stark

Pour l'effet Stark, il suffit ici d'examiner le cas le plus simple où seules les interactions électrostatiques sont considérées (on laisse de côté les couplages magnétiques du genre spin-orbite et donc toutes les complications liées à la structure fine). Dans ces conditions, le Hamiltonien H_0 de l'atome d'hydrogène en l'absence de champ appliqué est :

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r} \ . \tag{23.168}$$

Les états non-perturbés (sans spin) sont notés $|n, l, m\rangle$, avec la signification habituelle pour les nombres n, l, m : $\langle\vec{r}\,|n, l, m\rangle = R_{nl}(r)\,Y_{lm}(\theta,\,\phi)$. L'énergie à l'ordre zéro est donnée en (19.156), et est voisine de $-13,6/n^2\,\text{eV}$. En choisissant l'axe Oz le long du champ de module $\mathcal{E}$, la perturbation est :

$$V = -\,e\mathcal{E}z \tag{23.169}$$

puisque la force électrique agissant sur l'électron est $\vec{F} = -\vec{\nabla}V = e\vec{\mathcal{E}}$.

Avant de commencer les calculs, examinons brièvement les conditions (intuitives) de validité de la méthode de perturbation. On doit avoir :

$$\frac{|\langle n, l, m|V|n', l', m'\rangle|}{|E_n^{(0)} - E_{n'}^{(0)}|} \ll 1 \quad \Longleftrightarrow \quad |\langle n, l, m|e\mathcal{E}z|n', l', m'\rangle| \ll |E_n^{(0)} - E_{n'}^{(0)}| \ . \tag{23.170}$$

[46] C'est pour traiter l'effet Stark que Schrödinger [185] a introduit sa méthode de perturbation, dans son troisième article de 1926.

L'ordre de grandeur du premier membre est $|e\mathcal{E}a_0|$; le second est de l'ordre de e'^2/a_0. La condition (23.170) se récrit comme suit :

$$\mathcal{E} \ll \frac{|e|}{4\pi\varepsilon_0 a_0^2} \equiv \mathcal{E}_{\text{atomique}} \sim 5 \times 10^{11}\,\text{V/m} \ . \tag{23.171}$$

Le sens de cette dernière condition est évident : pour que l'on puisse parler de perturbation, il est nécessaire que le champ électrique appliqué de l'extérieur soit très petit devant le champ intra-atomique ; ce dernier étant gigantesque comparé aux champs du laboratoire, la condition *nécessaire* d'application de la méthode de perturbation est visiblement satisfaite.

Avant de mettre en œuvre la théorie des perturbations, les considérations de symétrie sont toujours utiles. H_0 est à symétrie sphérique et invariant par renversement du temps. Si V possède cette dernière symétrie, en revanche le couplage avec le champ électrique brise la symétrie sphérique : seule subsiste une symétrie de révolution autour Oz, et $\vec{L}^2$ n'est plus une constante du mouvement. En outre, V est invariant dans toute réflexion contenant l'axe Oz ; dans cette opération, L_z change de signe[47]. Les états propres $|n,l,m\rangle$ communs[48] à $(H,\ \vec{L}^2,\ L_z)$ ne différant que par le signe de la valeur propre associée à L_z auront donc la même énergie : l'énergie en présence du champ est donc fonction de n et de $|m|$, permettant de prévoir que la levée de dégénérescence ne sera que partielle. Enfin, V est changé en son opposé par parité[49] :

$$V' \stackrel{\text{déf}}{=} \Pi V \Pi^\dagger = -V \ ; \tag{23.172}$$

comme $\Pi\,|n,l,m\rangle = (-1)^l\,|n,l,m\rangle$ (voir (18.80)), l'élément de matrice $\langle n,l,m|z|n',l',m'\rangle$ est multiplié par $(-1)^{l+l'+1}$ par parité ; il est donc nul si l et l' sont de même parité :

$$\langle n,l,m|z|n',l',m'\rangle = 0 \quad \text{si } l \text{ et } l' \text{ de même parité} \ . \tag{23.173}$$

Examinons maintenant quantitativement l'effet du champ sur les premiers niveaux. En ce qui concerne le fondamental (non-dégénéré), $|1,0,0\rangle$, il n'y a pas de correction au premier ordre puisque $\langle 1,0,0|V|1,0,0\rangle = 0$. D'après (23.106) ou (23.108), la correction du second ordre est[50], *a priori* :

$$\varepsilon_{\text{fond}\,2} = -\sum_{n=2}^{+\infty}\sum_{l=0}^{+\infty}\sum_{m=-l}^{+l} \frac{|\langle 1,0,0|-e\mathcal{E}z|n,l,m\rangle|^2}{E_n^{(0)} - E_1^{(0)}} < 0 \ , \tag{23.174}$$

[47] tout comme la composante L_X située dans ce plan, cependant que la composante perpendiculaire au plan, L_Y, est inaltérée.

[48] De toute évidence, $\vec{L}^2$ est encore une constante du mouvement.

[49] Le champ extérieur n'appartient pas au système quantique et se factorise hors de tous les éléments de matrice.

[50] Ceci constitue le résultat élémentaire : on fait comme si le spectre ne contenait que des états liés. En toute rigueur, la sommation discrète doit être complétée par une intégrale sommant sur tous les états non-liés d'énergie positive – voir un exemple dans le problème 23.4.10.

où les $E_n^{(0)}$ sont données par (19.156). En fait, les éléments de matrice l pair et $m \neq 0$ sont nuls par symétrie, d'où :

$$\varepsilon_{\text{fond}\,2} = -\sum_{n=2}^{+\infty} \sum_{l\,\text{impair}}^{+\infty} \frac{|\langle 1,0,0| - e\mathcal{E}z|n,l,0\rangle|^2}{E_n^{(0)} - E_1^{(0)}} < 0 \ . \tag{23.175}$$

Grâce à une astuce très ingénieuse [186], il est possible de sommer cette série ([53], p. 266 et problème 23.4.8) ; on trouve :

$$\boxed{\varepsilon_{\text{fond}\,2} = -\frac{1}{2}\,\alpha\,\mathcal{E}^2 \qquad \alpha = 18\pi\varepsilon_0 a_0^3} \tag{23.176}$$

L'écriture $-(\alpha/2)\mathcal{E}^2$ est choisie pour rappeler qu'il s'agit d'un effet de polarisabilité (α désigne ponctuellement la polarisabilité – notation traditionnelle), la correction d'énergie quadratique en champ devant pouvoir s'écrire $-\int_0^{\mathcal{E}} \alpha\,\mathcal{E}'\mathrm{d}\mathcal{E}'$. Numériquement, la correction d'énergie due à la polarisabilité est extrêmement petite : avec $\mathcal{E} = 10^4$ V/cm, on trouve $|\varepsilon_{\text{fond}\,2}| \simeq 2 \times 10^{-10}$ eV $\simeq 56$ kHz.

Considérons maintenant le premier groupe d'états excités $n = 2$, définissant un sous-espace dégénéré de dimension égale à 4. Il convient de mettre en œuvre la technique exposée dans la sous-section 23.2.2. Notons que ce sous-espace est couplé au fondamental, puisque l'élément de matrice $\langle 1,0,0|z|2,1,0\rangle$ est différent de zéro ; comme on le sait (voir (23.128) et la fig. 23.1), la méthode de perturbation néglige ce type de couplage et projette V dans le sous-espace $n = 2$. Il faut donc calculer les éléments d'une matrice 4×4 ; en fait, compte tenu des considérations de symétrie ci-dessus, un seul élément de matrice est non-nul, c'est $\langle 2,0,0|V|2,1,0\rangle$. Un calcul simple d'intégrale donne :

$$\langle 2,0,0|V|2,1,0\rangle = 3e\mathcal{E}a_0 \stackrel{\text{déf}}{=} -v < 0 \ , \tag{23.177}$$

le signe venant de la charge de l'électron e ; les valeurs propres de l'opérateur projeté sont visiblement $\varepsilon_{2\,1} = 0, \pm v$; les vecteurs propres s'en déduisent :

$$|+v\rangle = \frac{1}{\sqrt{2}}\,(|2,0,0\rangle + |2,1,0\rangle) \ , \qquad |-v\rangle = \frac{1}{\sqrt{2}}\,(-|2,0,0\rangle + |2,1,0\rangle) \ . \tag{23.178}$$

$|\pm v\rangle$ a pour énergie $\pm v$ (l'état $|+v\rangle$ est au-dessus de l'état $|-v\rangle$). Ces deux états correspondent à la levée de dégénérescence, et sont les bons états propres (à cet ordre) dès que le champ est présent, aussi petit soit-il (c'est à nouveau un exemple de brisure de symétrie) ; pour $\mathcal{E} = 10^4$ V/cm, $v \simeq 1{,}6 \times 10^{-4}$ eV. Deux autres états restent dégénérés en gardant, à cet ordre, l'énergie $E_2^{(0)}$, ce sont les $|2,1,\pm 1\rangle$.

En utilisant la notation spectroscopique, les vecteurs propres (23.178) s'écrivent :

$$|+v\rangle = \frac{1}{\sqrt{2}}\,(|2S\rangle + |2P_0\rangle \ , \qquad |-v\rangle = \frac{1}{\sqrt{2}}\,(-|2S\rangle + |2P_0\rangle \ . \tag{23.179}$$

Ainsi, le champ électrique lève partiellement la dégénérescence "accidentelle" du champ coulombien, et mélange deux états de symétrie différente vis-à-vis du moment cinétique

orbital[51]. Ceci a des conséquences importantes, en particulier pour la durée de vie de certains états excités.

En particulier, l'état $|2S\rangle$ a une durée de vie très longue, voisine de la seconde[52], au contraire de l'état $|2P\rangle$ dont la durée de vie est de l'ordre de 10^{-9} s. La raie la plus "rouge" de la série de Lyman[53], appelée L_α, correspond précisément à la transition $|2P\rangle \rightarrow |1S\rangle$. Le calcul perturbatif ci-dessus montre que le champ électrique, parce qu'il contamine l'état 2S par l'état 2P, a un effet spectaculaire sur la durée de vie de l'état métastable : tout champ électrique, même petit, est donc capable de détruire le caractère métastable. Si l'atome est préparé dans l'état 2S et qu'un champ électrique est présent, volontairement ou non, la métastabilité est anéantie et l'atome peut rayonner presque normalement grâce à sa composante P. Ce phénomène porte le nom de *quenching* de métastabilité.

♦ *Remarques*

1. Le calcul qui vient d'être fait est tout à fait recevable, dans le sens où il rend très bien compte, qualitativement et quantitativement, de l'effet Stark tel qu'il est observé. Il est utile de remarquer que cette adéquation fait fi d'une "irrégularité" commise depuis le début : la perturbation est ici singulière et l'hypothèse de l'existence d'une série de perturbation convergente est prise en défaut. On peut s'en convaincre sur des arguments purement physiques, esquissés ci-dessous.

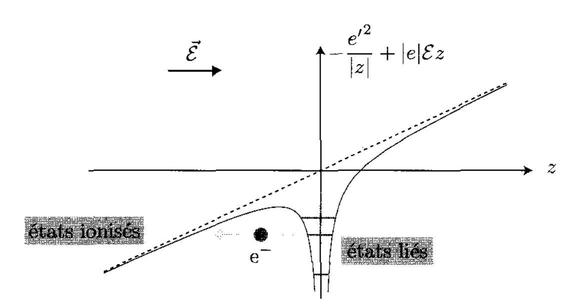


Figure 23.4: Variation de l'énergie potentielle totale V_{tot} le long de l'axe du champ, montrant qu'il existe des états ionisés de même énergie que les états liés. Avec les champs usuels, la droite est presque horizontale ; d'un point de vue pragmatique, c'est ce qui justifie l'emploi de la théorie de perturbation.

Dès que l'atome est soumis à un champ électrique, si faible soit-il, il n'y a plus d'états liés au sens strict du terme. En effet, l'énergie potentielle totale de l'électron

[51] Un tel mélange d'états s'appelle traditionnellement *hybridation*, terme que l'on retrouve également en Chimie à propos des orbitales atomiques ou moléculaires.

[52] Ceci tient au fait que, compte tenu de la symétrie orbitale des deux états initial et final, la transition 2S $\rightarrow$ 1S se fait par un processus à deux photons dont la probabilité est très faible. Cette durée de vie justifie que l'on déclare l'état 2S *métastable*.

[53] tout entière dans l'ultraviolet (voir Tome I, p. 164).

est V_{tot} :

$$V_{\text{tot}} = -\frac{e'^2}{r} - e\mathcal{E}z = -\frac{e'^2}{r} + |e|\mathcal{E}z \ . \tag{23.180}$$

En supposant le champ $\vec{\mathcal{E}}$ dirigé dans le sens des z positifs, et compte tenu du signe de la charge e, on a :

$$\lim_{z\to+\infty} V_{\text{tot}} = +\infty \ , \qquad \lim_{z\to-\infty} V_{\text{tot}} = -\infty \tag{23.181}$$

avec toujours $\lim_{|z|\to 0} V_{\text{tot}} = -\infty$. Un potentiel ayant de tels comportements n'a pas d'états liés ; ceci traduit la possibilité pour l'électron de passer par effet-tunnel de n'importe quel état localisé à un état dissocié (ionisé). Dit autrement, supposons l'électron préparé, en l'absence de champ, dans l'état fondamental ; cet état est stable, son énergie (négative) est inférieure à la valeur limite du potentiel Coulombien à l'infini, l'électron est piégé et reste lié. En revanche, dès que l'on branche le champ, le même électron voit des états de *continuum* accessibles à énergie constante : sa probabilité de passage tunnel devient finie, quoique petite. C'est l'expression quantique d'un phénomène banal : un champ électrique peut ioniser un atome.

Bien sûr, toute la question est de savoir combien vaut cette probabilité d'ionisation. Plus le champ est faible, plus la barrière-tunnel est épaisse ; la probabilité de passage par effet-tunnel ayant une dépendance exponentielle par rapport à l'épaisseur[54], elle est très petite pour les champs usuels, donnant un temps moyen d'ionisation exponentiellement grand. Sans devoir mettre des nombres plus ou moins précis dans le problème[55], le calcul perturbatif est validé par le fait que le temps d'ionisation est immensément long avec les champs ordinaires, tant que l'atome n'est pas fortement excité. Cette conclusion doit être reconsidérée pour les champs intenses que l'on peut rencontrer dans les milieux astrophysiques, mais aussi pour les états très excités impliqués, notamment, dans les atomes de Rydberg.

2. La discussion de la dégénérescence particulière du champ Coulombien (chapitre 19, sous-section 19.2.3) a donné l'occasion d'évoquer les cordonnées paraboliques, où les variables se séparent également [53]. Epstein [187] a donné le premier traitement complet de l'effet Stark dans cette représentation. ♦

23.3.2 Effet Zeeman

À nouveau, on choisit l'axe Oz le long du champ, ici magnétique et noté $\vec{\mathcal{B}}$. On conçoit qu'une description cohérente exige de prendre en compte tous les effets magnétiques

[54]d'où l'extraordinaire sensibilité du microscope à effet-tunnel.

[55]Avec un champ de 10^4 V/m, l'épaisseur L de la barrière à l'énergie $E_1 \simeq -13,6$ eV est à peu près égale à 3 mm, une longueur astronomique à l'échelle atomique. Si on admet une dépendance exponentielle de la probabilité (voir Tome I, section 15.5), elle ne peut être que de la forme $\mathrm{e}^{-L/a}$, où l'autre longueur a est forcément atomique, ce qui donne une probabilité de l'ordre de $\exp(-3 \times 10^7)$, autant dire zéro.

simultanément et c'est pourquoi le Hamiltonien H_0 doit ici contenir le couplage spin-orbite, responsable de la structure fine. Ceci étant, H_0 s'écrit :

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r} + A(r)\,\vec{L}.\vec{S} \ , \tag{23.182}$$

où la fonction $A(r)$ est donnée en (20.20). En présence du champ $\vec{\mathcal{B}}$, et en négligeant le terme diamagnétique, ce Hamiltonien doit être complété[56] par V :

$$V = -\frac{e}{2m}(L_z + g_e S_z)\mathcal{B} \qquad (g_e = 2,002\,3\ldots \simeq 2) \ , \tag{23.183}$$

de sorte que le Hamiltonien en présence du champ magnétique est :

$$H = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r} + A(r)\vec{L}.\vec{S} - \frac{e}{2m}(L_z + g_e S_z)\mathcal{B} \ ; \tag{23.184}$$

le terme de structure fine étant explicitement introduit, on n'est plus tenu[57] de considérer seulement la situation où le champ est assez intense pour qu'elle puisse être négligée.

Si le champ est très faible, le couplage V est très petit devant le terme spin-orbite, ce dernier étant lui-même très inférieur aux deux premiers termes de H_0 ; c'est cette situation qui va être traitée et qui, strictement, s'appelle l'effet Zeeman[58]. Ce choix étant fait, les états propres à l'ordre zéro sont ceux de H_0 et se notent $|n, j, m_j, l, s\rangle$; ils sont communs à (H_0, $\vec{J}^2$, J_z, $\vec{L}^{\,2}$, $\vec{S}^{\,2}$). Hors dégénérescence accidentelle, chaque niveau est dégénéré $2j+1$ fois ($m_j = -j, -j+1, \ldots +j$) et c'est à nouveau la théorie pour un niveau dégénéré qu'il faut appliquer : considérer exclusivement la matrice de V dans le sous-espace auquel on s'intéresse, oublier le reste et diagonaliser. Pour trouver la matrice de V dans ce sous-espace, deux méthodes peuvent être suivies :

- V s'exprime à l'aide de L_z et de S_z, donc sa matrice sur la base $\{|n, l, m_l, m_s\rangle\}$ se détermine facilement. Il reste ensuite à utiliser les relations entre cette base et la base $\{|n, j, m_j, l, s\rangle\}$ – qui impliquent les coefficients de Clebsch - Gordan – pour en déduire la matrice de V sur la base propre de H_0.
- On peut aussi invoquer le théorème de Wigner - Eckart, d'après lequel, à l'intérieur du sous-espace $\{|n, j, m_j, l, s\rangle\}$, $\vec{L} + g_e\vec{S}$ est un opérateur vectoriel proportionnel à $\vec{J}$, avec un facteur de proportionnalité bien déterminé (voir (18.260), et problème 18.7.6) :

$$\vec{L} + g_e\vec{S} = \frac{\langle(\vec{L} + g_e\vec{S}).\vec{J}\rangle}{\hbar^2 j(j+1)}\,\vec{J} \equiv g_{lsj}\,\vec{J} \ . \tag{23.185}$$

[56] L'ajout *à la main* du terme en S_z est justifié par la limite faiblement relativiste conduisant au Hamiltonien de Pauli (voir chapitre 20, section 20.7, et notamment (20.235))

[57] au contraire de la présentation de la section 20.2.

[58] L'autre cas où V_{magn} est petit devant le couplage avec $\vec{B}$ donne lieu à l'effet dit Paschen - Back. Il est bien clair que l'on passe continûment d'un cas à l'autre en augmentant graduellement le champ extérieur.

g_{lsj} désigne un nombre pur (appelé *facteur de Landé*), qui s'exprime à l'aide des nombres quantiques l, s et j ; en écrivant $\vec{L}.\vec{S} = \frac{1}{2}(\vec{J}^2 - \vec{L}^2 - \vec{S}^2)$, il vient :

$$(\vec{L} + g_e\vec{S})(\vec{L} + \vec{S}) = \frac{1}{2}\left[(g_e + 1)\vec{J}^2 - (g_e - 1)\vec{L}^2 + (g_e - 1)\vec{S}^2)\right] \ . \tag{23.186}$$

Si on choisit la vleur approchée $g_e = 2$, g_{lsj} est un rationnel simple[59] :

$$g_{lsj} \simeq 1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)} \tag{23.187}$$

Les égalités (23.183) et (23.185) montrent que la matrice de l'opérateur projeté V_{ap} (voir (23.128)) est déjà diagonale ; les corrections au premier ordre pour l'énergie se lisent sur la diagonale et sont donc :

$$\varepsilon_{nj\,1} = -\,g_{lsj}\,\frac{e\mathcal{B}}{2m}\,m_j\hbar \equiv -g_{lsj}m_j\mu_{\text{B}}\mathcal{B} \ . \tag{23.188}$$

Le champ magnétique lève complètement la dégénérescence et fait éclater chaque niveau atomique en $2j + 1$ niveaux équidistants, l'écart entre deux niveaux consécutifs étant proportionnel à l'intensité $\mathcal{B}$ du champ appliqué. Pour l'atome d'hydrogène avec son unique électron de spin 1/2, j est forcément demi-entier et chaque niveau engendre un nombre *pair* de sous-niveaux Zeeman.

Cette levée complète résulte du fait que le champ magnétique, au contraire du champ électrique, brise l'invariance par renversement du temps : deux états se distinguant par le signe de m_j ont des énergies opposées par rapport à l'énergie en champ nul. Le traitement précédent ne vaut que quand l'éclatement Zeeman reste petit devant la structure fine, puisque chaque multiplet a été traité indépendamment de tous les autres. Visiblement, certaines composantes Zeeman de deux multiplets adjacents ont tendance à se rapprocher lorsque le module du champ augmente (croisement de niveaux ?) ; quand le champ devient assez intense, tout doit être reconsidéré, afin de mettre en œuvre un traitement (toujours approximatif) plus astucieux (voir problème 23.4.11) ; ceci constitue un exemple classique de croisements de niveaux évités ou persistents.

23.4 Exercices et problèmes

23.4.1 Méthode variationnelle

1. Pour décrire l'état fondamental d'un puits infini à une dimension, on choisit l'ensemble de fonctions $\psi_\alpha(x)$:

$$\psi_\alpha(x) = C_\alpha\, x^\alpha (a - x)^\alpha \qquad (0 < x < a) \tag{23.189}$$

[59]Dans la limite des très grands nombres quantiques, s est négligeable devant l et $g_{lsj} \simeq 1$.

où l'exposant α, supposé positif, est le paramètre variationnel. Pour les calculs, on passera par l'intermédiaire de la fonction[60] $B(x, y)$ définie comme :

$$B(x,y) \stackrel{\text{déf}}{=} \int_0^1 u^{x-1}(1-u)^{y-1}\,\mathrm{d}u = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \ , \tag{23.190}$$

où Γ désigne la fonction Gamma d'Euler, satisfaisant la relation $\Gamma(x+1) = x\Gamma(x)$.

(a) Tracer qualitativement $\psi_\alpha(x)$ en distinguant les deux cas $0 < \alpha < 1$ et $\alpha > 1$.

(b) Exprimer la constante de normalisation C_α à l'aide de la fonction B.

(c) Calculer la valeur moyenne de l'énergie, $E(\alpha)$, et la mettre sous la forme :

$$E(\alpha) = \frac{\hbar^2}{ma^2}\, Q(\alpha) \ , \tag{23.191}$$

où $Q(\alpha)$ est une fraction rationnelle.

(d) Déduire de ce qui précède la valeur optimale de α, α_{opt}, et la valeur correspondante de l'énergie, E_{opt}. Comparer cette dernière à la valeur exacte.

2. Une particule de masse m est confinée dans le potentiel quartique $V(x)$:

$$V(x) = V_0 \left[2\left(\frac{x}{a}\right)^4 + g\left(\frac{x}{a}\right)^2\right] \qquad (g \in \mathbb{R}) \ . \tag{23.192}$$

(a) Tracer l'allure de $V(x)$ selon que g est positif ou négatif.

(b) g étant supposé positif, on choisit pour décrire l'état fondamental la gaussienne normalisée :

$$\psi(x) = \left(\frac{\alpha}{\pi a^2}\right)^{1/4} \mathrm{e}^{-\alpha x^2/(2a^2)} \ . \tag{23.193}$$

Ce choix serait-il judicieux dans le cas $g < 0$?

(c) On pose $gV_0 = m\omega^2 a^2/2$. Calculer la valeur moyenne de l'énergie $E(\alpha)$.

(d) Montrer que la meilleure valeur de α, α_{opt}, est donnée par une équation du troisième degré. Obtenir les développements de α_{opt} pour $g \ll 1$ et $g \gg 1$ et en déduire ceux de E_{opt}. Tracer qualitativement E_{opt} en fonction de g ; commenter.

3. Une particule de masse m se déplace dans le potentiel :

$$V(x) = V_0 \left[\left(\frac{x}{x_0}\right)^2 - 1\right]^2 \qquad (V_0 > 0,\, x_0 > 0) \ . \tag{23.194}$$

(a) Tracer l'allure de $V(x)$.

(b) On choisit une fonction d'essai sous la forme :

$$\psi(x) = c_1\psi(x-a) + c_2\psi(x+a) \equiv c_1\psi_1(x) + c_2\psi_2(x) \ . \tag{23.195}$$

où les coefficients c_i sont réels et où $\psi(x) = (\alpha/\pi)^{1/4}\mathrm{e}^{-\alpha x^2/2}$. Justifier ce choix et exprimer a et α en fonction de x_0 et V_0 (on posera $\varepsilon \stackrel{\text{déf}}{=} [8mx_0^2V_0/\hbar^2]^{1/2}$).

[60] $B = \beta$ majuscule (prononcer "béta"). $B(p, q)$ est la fonction d'Euler de première espèce.

(c) Écrire les équations variationnelles en fonction des intégrales :

$$J \stackrel{\text{déf}}{=} \langle\psi_1|H|\psi_1\rangle \ , \quad K \stackrel{\text{déf}}{=} \langle\psi_1|H|\psi_2\rangle \ , \quad S \stackrel{\text{déf}}{=} \langle\psi_1|\psi_2\rangle \ . \tag{23.196}$$

(d) Trouver les énergies et les fonctions propres approchées.

(e) Calculer effectivement les différentes intégrales. En déduire la différence d'énergie ΔE entre les deux états ainsi trouvés ; étudier sa variation en fonction de la hauteur de la barrière de potentiel.

4. Pour l'atome d'hydrogène dans $\mathbb{R}^D$, calculer l'expression $E(\lambda)$ de l'énergie pour la fonction d'essai $C\mathrm{e}^{-\lambda r/a_0}$. Tracer la variation de l'énergie optimisée E_{opt} en fonction de la dimension D.

5. Soit un atome alcalin (Li, Na, K, ...). La question est de rendre compte des effets d'écran et de la répulsion dus aux électrons internes en modifiant le potentiel effectif ressenti par l'électron célibataire.

(a) Un premier choix consiste à prendre l'énergie potentielle de l'électron périphérique sous la forme :

$$V(r) = -\frac{Ze'^2}{r} + \frac{A}{r^2} \ . \tag{23.197}$$

i. Trouver r_0 où V est minimum et exprimer A en fonction de e', r_0 et Z.

ii. La fonction variationnelle est choisie comme suit :

$$\psi_\alpha(r) = C\, r\mathrm{e}^{-\alpha r/a_0} \ . \tag{23.198}$$

Trouver la valeur optimale de α, et en déduire l'énergie E_{opt}.

(b) On choisit maintenant un potentiel de Yukawa[61] :

$$V(r) = -\frac{a}{r}V_0\mathrm{e}^{-r/a} \qquad (a > 0,\, V_0 > 0) \ . \tag{23.199}$$

i. Avec la fonction d'essai $\psi_\beta(r) = C\mathrm{e}^{-\beta r/a_0}$, calculer la valeur moyenne de l'énergie potentielle, et la mettre sous la forme $E = |E_1| f(\beta\,;\, \varepsilon,\, a_0/a)$, où $\varepsilon \doteq 2maa_0V_0/\hbar^2$.

ii. Préciser le comportement de $|E|$ pour $\beta \ll 1$ et $\beta \gg 1$. En déduire que si ε est inférieur à une certaine valeur ε_c, il n'existe pas d'état lié de ce type[62].

iii. Par un processus de limite approprié, retrouver la valeur (exacte) de l'énergie du fondamental de l'hydrogène. Pourquoi retrouve-t-on de fait l'énergie *exacte* ?

6. On reconsidère le double puits de potentiel $V(x)$, exercice 2, éq. (23.192), avec $g < 0$. Les minima de potentiel se trouvent aux points d'abscisses $x = \pm a/2$. Préciser la signification physique de V_0. On pose $V_0 = \frac{1}{32}m\omega^2a^2$; quel est le sens physique de la pulsation ω ?

[61] Les états liés sphériques d'un tel potentiel sont traités exactement dans le problème 19.3.9.

[62] Le calcul détaillé de ε_c n'est pas nécessaire pour une argumentation précise.

(a) Dans toute la suite $\psi(x)$ désigne une fonction positive "en cloche" bien localisée au voisinage de $x = 0$, paire et normalisée à l'unité. On cherche maintenant à approcher les états[63] propres du Hamiltonien en choisissant des fonctions d'essai de la forme :

$$\Psi(x) = \cos\theta\,\psi_{\rm d}(x) + \sin\theta\,\psi_{\rm g}(x) \tag{23.200}$$

où $\psi_{\rm g}(x) = \psi(x + a/2)$, $\psi_{\rm d}(x) = \psi(x - a/2)$; le paramètre variationnel est ici l'angle θ. On pose :

$$E_0 = \int_{-\infty}^{+\infty} \psi_{\rm d}(x) H \psi_{\rm d}(x)\,{\rm d}x = \int_{-\infty}^{+\infty} \psi_{\rm g}(x) H \psi_{\rm g}(x)\,{\rm d}x \ , \tag{23.201}$$

$$\hbar\Omega = -\int_{-\infty}^{+\infty} \psi_{\rm g}(x) H \psi_{\rm d}(x)\,{\rm d}x \ . \tag{23.202}$$

Calculer l'expression de l'énergie E en fonction de E_0, Ω et θ.

(b) Lorsque la hauteur de la barrière est grande[64] par rapport à l'énergie de point zéro au fond de chaque cuvette, le paramètre Ω est positif. En déduire par variation les énergies des deux premiers états du double puits (fondamental et premier état excité) et l'expression des fonctions propres correspondantes (on négligera le produit scalaire $\langle\psi_{\rm g}|\psi_{\rm d}\rangle$). Dessiner celles-ci.

(c) On suppose que l'état initial de la particule est$\Psi(x, t = 0) = \psi_{\rm d}(x)$. Montrer que la valeur moyenne de la coordonnée à l'instant t est :

$$\langle x\rangle(t) = \frac{a}{2}\cos 2\Omega t \ . \tag{23.203}$$

et en déduire le sens physique de Ω.

(d) Calculer explicitement Ω en prenant :

$$\psi(x) = \left(\frac{\alpha}{\pi}\right)^{1/4} {\rm e}^{-\alpha x^2/2} \qquad (\alpha = m\omega/\hbar) \tag{23.204}$$

Vérifier l'affirmation ci-dessus sur le signe de Ω. Supposant $\hbar^{-1}m\omega a^2$ très grand devant 1, montrer que la fréquence-tunnel est approximativement donnée par $\Omega \simeq a\hbar^{-1}V_0\,{\rm e}^{-b\frac{V_0}{\hbar\omega}}$, où a et b sont des nombres.

7. On applique la méthode variationnelle dans le cas de l'atome d'hydrogène soumis à un champ électrique $\vec{\mathcal{E}}$ parallèle à Oz, en choisissant comme fonction d'essai :

$$\psi_\lambda(r) = C(1 + \lambda z){\rm e}^{-r/a_0} \ , \tag{23.205}$$

où C est la constante de normalisation et λ le paramètre variationnel.

(a) Justifier physiquement ce choix, et calculer C.

(b) Trouver l'expression de l'énergie en fonction de λ.

(c) Que donne la minimisation ? Commenter.

---

[63] À une dimension, les états liés peuvent toujours être pris réels.

[64] Cette hypothèse est nécessaire pour que l'emploi de la méthode variationnelle ait ici un sens.

23.4.2 Champ auto-cohérent à une dimension : deux fermions en interaction de contact

Deux fermions identiques de masse m sont en interaction de contact entre eux et avec un centre (le "noyau") de masse infinie [188]. Tout se passant à une dimension, le Hamiltonien s'écrit :

$$H = -\frac{\hbar^2}{2m}\Big(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\Big) - Ze'^2[\delta(x_1) + \delta(x_2)] + e'^2\delta(x_1 - x_2) \ . \qquad (23.206)$$

Il s'agit de traiter l'état fondamental par la méthode SCF.

1. Expliquer pourquoi il est suffisant de considérer la seule fonction d'espace :
$$\Phi(x_1, x_2) \stackrel{\text{déf}}{=} \phi(x_1)\phi(x_2) \ . \qquad (23.207)$$

2. En écrivant la stationnarité de la fonctionnelle $E[\phi] \stackrel{\text{déf}}{=} \langle\phi|H|\phi\rangle - E\langle\phi|\phi\rangle$, montrer que l'équation SCF s'écrit :
$$-\frac{\hbar^2}{2m}\frac{\partial^2\phi}{\partial x^2} - Ze'^2\delta(x) + e'^2[\phi(x)]^3 = \varepsilon\phi(x) \ ; \qquad (23.208)$$
interpréter le terme cubique et identifier le Hamiltonien H^{SCF}. Dans la suite, on note $\gamma = 2me'^2/\hbar^2 \equiv 2/a_0$, $\lambda = Z\gamma$ et $k^2 = -2mE/\hbar^2$ ($k > 0$).

3. Y-a-t-il ici un terme d'échange ?

4. Montrer que :
$$E^{\text{SCF}} \stackrel{\text{déf}}{=} \langle\Phi|H|\Phi\rangle = 2\varepsilon - e'^2\int_{\mathbb{R}}[\phi(x)]^4\,\mathrm{d}x \ ; \qquad (23.209)$$
Pourquoi cette valeur moyenne n'est-elle pas simplement égale à 2ε ?

5. Calculer le saut de la dérivée $\phi'(x)$ en $x = 0$.

6. Montrer que, par une intégration, (23.208) conduit à :
$$\Big(\frac{\mathrm{d}\phi}{\mathrm{d}x}\Big)^2 = \frac{\gamma}{2}[\phi(x)]^4 + k^2[\phi(x)]^2 + C \ ; \qquad (23.210)$$
combien vaut la constante C ?

7. En déduire que la fonction $\phi(x)$ est :
$$\phi(x) = 2Ak\sqrt{\frac{2}{\gamma}}\frac{\mathrm{e}^{-k|x|}}{1 - A^2\mathrm{e}^{-2k|x|}} \ , \qquad (23.211)$$
où A est la constante d'intégration.

8. Écrire la relation traduisant le résultat obtenu en 5, et en déduire A^2. Comment se situe A par rapport à 1 ? Commenter.

9. Montrer que $k = (Z - \frac{1}{2})/a_0$.

10. Quelle est la condition pour qu'il existe une solution liée ? Commenter.

11. Trouver l'expression de l'énergie E^{SCF} en fonction de Z, et commenter sa variation.

23.4.3 La méthode de Brillouin - Wigner

Dans la suite, $|\Psi_0\rangle$ désigne l'état fondamental d'un Hamiltonen H_0, d'énergie E_0. Il s'agit de trouver l'état fondamental $|\Psi\rangle$ de $H \stackrel{\text{déf}}{=} H_0 + V$; la norme de $|\Psi\rangle$ est ajustée de sorte que $\langle\Psi_0|\Psi\rangle = 1$.

1. Montrer que le déplacement d'énergie dû à V est donné par :

$$E - E_0 = \langle\Psi_0|V|\Psi\rangle \ . \tag{23.212}$$

2. Soit le projecteur $P \stackrel{\text{déf}}{=} |\Psi\rangle\langle\Psi|$, et son complémentaire $Q = \mathbf{1} - P$. Montrer que :

$$(z\mathbf{1} - H_0)|\Psi\rangle = (z\mathbf{1} - E\mathbf{1} + V)|\Psi\rangle \ . \tag{23.213}$$

3. En déduire l'égalité :

$$|\Psi\rangle = |\Psi_0\rangle + QG_0(z)(z\mathbf{1} - E\mathbf{1} + V)|\Psi\rangle \ , \tag{23.214}$$

   où $G_0(z)$ est la résolvante associée à H_0.

4. En itérant l'égalité précédente, montrer que :

$$|\Psi\rangle = \sum_{n=0}^{+\infty} Q[G_0(z)(z\mathbf{1} - E\mathbf{1} + V)]^n|\Psi_0\rangle \ , \tag{23.215}$$

5. En déduire que l'énergie est donnée par :

$$E = E_0 + \sum_{n=0}^{+\infty}\langle\Psi_0|V[QG_0(z)(z\mathbf{1} - E\mathbf{1} + V)]^n|\Psi_0\rangle \ , \tag{23.216}$$

6. Ces résultats acquis, les formules de Brillouin - Wigner s'obtiennent immédiatement en choisissant $z = E$:

$$|\Psi\rangle = \sum_{n=0}^{+\infty} Q[G_0(E)V]^n|\Psi_0\rangle \ , \qquad E = E_0+\sum_{n=0}^{+\infty}\langle\Psi_0|V[QG_0(E)V]^n|\Psi_0\rangle \ . \tag{23.217}$$

7. Comparer ces expressions à leurs équivalents obtenus dans la méthode de Rayleigh - Schrödinger.

23.4.4 Exemples simples de perturbations

Pour chacun des cas suivants, traiter par perturbation au premier ordre les états dégénérés et jusqu'au second ordre inclus celui qui ne l'est pas :

1. $H_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & E_0 \end{bmatrix}$, $V = \begin{bmatrix} 0 & v & 0 \\ v & 0 & v \\ 0 & v & 0 \end{bmatrix}$

2. $H_0 = \begin{bmatrix} E_0 & 0 & 0 \\ 0 & E_0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $V = \begin{bmatrix} 0 & v & v \\ v & 0 & 0 \\ v & 0 & 0 \end{bmatrix}$

23.4.5 Deux oscillateurs couplés

Deux oscillateurs harmoniques identiques (m, ω) sont couplés par un terme bilinéaire $gm\omega^2 x_1 x_2$. Trouver la correction d'énergie au second ordre pour l'état fondamental et les levées de dégénérescence des deux premiers états excités. Comparer à la solution exacte.

23.4.6 Effet anharmonique pour un oscillateur

Afin de rendre compte de l'asymétrie d'un fond de potentiel pour une particule de masse m, on définit l'approximation anharmonique suivante :

$$V(x) = \frac{1}{2}m\omega^2 x^2 \left(1 + 4g\frac{x}{x_0}\right) \ , \qquad x_0 = \sqrt{\frac{\hbar}{2m\omega}} \ . \tag{23.218}$$

Trouver la correction d'énergie pour le n^{e} niveau. Que pensez-vous de l'emploi de la méthode perturbative dans cette situation ?

23.4.7 Terme de contact pour l'électron dans l'atome d'hydrogène

Outre le champ coulombien du noyau, on admet que l'électron est soumis au potentiel de contact :

$$V_{\text{c}}(x) = g\varepsilon a_0^3 \delta(\vec{r}) \ , \quad \delta(\vec{r}) = \frac{1}{4\pi r^2}\,\delta(r) \ , \quad \varepsilon = \frac{e'^2}{a_0} \ . \tag{23.219}$$

1. Pour l'état fondamental, calculer les corrections d'énergie au premier et au second ordre, $\Delta E^{(1)}$ et $\Delta E^{(2)}$.

   On donne les valeurs à l'origine des fonctions propres s de l'atome d'hydrogène :

   - états liés : $\psi_{n00}(0) = (\pi a_0^3 n^3)^{-1/2}$
   - états non-liés : $\psi_{E00}(0) = (\pi a_0^3 E)^{-1/2}\left[1 - \mathrm{e}^{-\pi\sqrt{2\varepsilon/E}}\right]^{-1/2}$

2. On utilise maintenant le Principe variationnel.

(a) Si $g > 0$, montrer que $E_1(g) > E_1(0)$

(b) Dans le cas $g < 0$, on choisit comme fonction d'essai $e^{-\lambda r/a_0}$. Montrer que l'énergie de l'état fondamental n'est pas bornée.

23.4.8 Effet Stark pour l'hydrogène

On soumet un atome d'hydrogène à un champ électrique constant et homogène de module $\mathcal{E}$; l'axe de quantification Oz est choisi le long du champ. Le Hamiltonien est $H = H_0+V$ où le terme complémentaire $V = -e\mathcal{E}z$ est traité comme une perturbation. Les états propres[65] de H_0 sont notés $|\psi_{nlm}\rangle$ et ont l'énergie E_n.

1. Estimer l'ordre de grandeur du champ électrostatique intra-atomique.

2. (a) Montrer que $\langle\psi_{nlm}|V|\psi_{n'l'm'}\rangle = 0$ si l et l' sont de même parité .

   (b) En utilisant la réflexion-miroir par rapport à un plan contenant Oz, montrer que les deux états $\psi_{nl\,\pm m}$ conservent la même énergie en présence du champ. A quelle autre symétrie est associée cette dégénérescence ?

   (c) Montrer que $\langle\psi_{nlm}|V|\psi_{n'l'm'}\rangle = 0$ si $m \neq m'$.

3. Compte tenu de ce qui précède, il n'y a pas de correction linéaire en champ pour l'état fondamental. Au plus bas ordre d'approximation, on attend donc une variation quadratique en $\mathcal{E}$; traditionnellement, on pose :

$$\Delta E_{100} = -\frac{1}{2}\,\alpha\,\mathcal{E}^2 \ . \qquad (23.220)$$

   (a) Quel est le signe de la grandeur α, appelée polarisabilité ? Quelle est sa dimension ? Quel est son sens physique ?

   (b) Écrire l'expression formelle de α.

   (c) Pour sommer la série donnant α, on utilise la méthode de Dalgarno et Lewis [186], en cherchant un opérateur F tel que $[F, H_0]\,|\psi_{100}\rangle = z|\psi_{100}\rangle$. Exprimer α sous la forme d'un élément de matrice.

   (d) On suppose que F ne dépend que de la coordonnée $\vec{r}$. Etablir une équation différentielle pour $F(\vec{r})$ et en chercher une solution sous la forme :

$$F(\vec{r}) = f(r)\,Y_{lm}(\theta,\phi) \ . \qquad (23.221)$$

   Préciser l et m ; vérifier qu'il existe une solution polynômiale :

$$f(\vec{r}) = c_0 + c_1 r + c_2 r^2 \ , \qquad (23.222)$$

   et déterminer les coefficients c_i.

[65] Les premières fonctions radiales sont données en (19.187) et (19.188).

(e) Obtenir finalement l'expression de la polarisabilité α en fonction de a_0 et l'expression explicite de la fonction d'onde fondamentale perturbée au premier ordre ; interpréter ce résultat.

4. On s'intéresse maintenant aux corrections linéaires pour le niveau dégénéré $n = 2$.

(a) Préciser sans calcul les éléments de matrice $\langle\psi_{2lm}|V|\psi_{2l'm'}\rangle$ qui sont nuls.

(b) En déduire que les deux états $|\psi_{21\,\pm 1}\rangle$ ont une correction d'énergie nulle au premier ordre.

(c) En désignant par v l'élément de matrice $\langle\psi_{210}|V|\psi_{200}\rangle$, exprimer en fonction de v les corrections d'énergie et en déduire, à cette approximation, les nouveaux états propres en présence du champ.

(d) Trouver l'expression de v sachant que ($\rho = r/a_0$) :

$$\psi_{200} = \frac{1}{\sqrt{4\pi}}\,(2a_0)^{-3/2}(2-\rho)\,\mathrm{e}^{-\rho/2} \qquad \text{(état 2S)} \ , \tag{23.223}$$

$$\psi_{210} = \frac{1}{\sqrt{4\pi}}\cos\theta\,(2a_0)^{-3/2}\rho\,\mathrm{e}^{-\rho/2} \qquad \text{(état 2P, } m=0\text{)} \ . \tag{23.224}$$

Calculer v numériquement pour $\mathcal{E} = 10^4$ V/cm. Reprendre ce qui précède pour le niveau $n = 3$.

23.4.9 Effet de taille finie du noyau

Pour un atome lourd, on peut admettre, à une bonne approximation, qu'un électron de la couche K est soumis esentiellement au seul champ du noyau de charge $Z|e|$. Pour tenir compte de l'effet de taille finie du noyau, de rayon[66] r_0, on admet que le champ du noyau est celui d'une distribution de charge positive continue répartie uniformément dans la sphère de rayon r_0.

1. Trouver l'énergie potentielle de l'électron K, $V(r)$, $\forall r$.

2. $V_\mathrm{C}(r)$ désignant le terme coulombien pur, on pose :

$$W(r) = V(r) - V_\mathrm{C}(r) \ , \tag{23.225}$$

et on traite $W(r)$ par perturbation. Écrire l'expression de $W(r)$.

c) Expliquer le mode de calcul au premier ordre pour un niveau n quelconque. Montrer que :

$$\langle\psi_{nlm}|W|\psi_{n'l'm'}\rangle = \delta_{ll'}\,\delta_{mm'}\int_0^{r_0} r^2\mathrm{d}r\,\left[R_{nl}(r)\right]^2\,W(r) \ . \tag{23.226}$$

3. L'échelle typique de variation de R_{nl} est na_0. En utilisant le fait que $r_0 \ll a_0$, calculer approximativement l'intégrale radiale et en déduire la correction d'énergie :

$$\Delta E_{nl} = \delta_{l0}\,\frac{2\pi}{5}\,Ze'^2r_0^2\,|\psi_{n00}(0)|^2 \ . \tag{23.227}$$

[66] r_0 est de l'ordre de quelques fermis (1 F = 10^{-15} m). Par ailleurs, on rappelle que $Y_{00} = (4\pi)^{-1/2}$.

23.4.10 Rôle du continuum pour la correction du second ordre

En toute rigueur, la formule donnant la correction d'énergie au second ordre doit également contenir une somme sur les états non-liés, quand ils existent. Cet exercice, fort académique comme on pourra en juger, montre explicitement l'erreur qui peut être commise quand on oublie ce fait [189].

1. Pour une particule de masse m, trouver tous les états, liés et non-liés, d'un puits de Dirac attractif $-g\delta(x)$ à une dimension ($g > 0$).
2. On ajoute à ce puits la perturbation $V(x) = \lambda\delta(x)$, avec $|\lambda| \ll g$.

   Calculer les corrections d'énergie au premier et second ordre pour l'état lié. Comparer à la solution exacte et commenter.

23.4.11 Effet Zeeman en champ assez fort : croisements de niveaux ?

Il s'agit de traiter en détail l'action d'un champ magnétique sur les premiers niveaux de l'atome d'hydrogène et, en particulier, d'examiner l'évolution des niveaux dont les énergies se rapprochent les unes des autres quand le module du champ augmente à partir de zéro[67]. On prendra partout $g_\mathrm{e} \simeq 2$.

1. En présence d'interaction spin-orbite, le Hamiltonien est $H = H_0 + A(r)\,\vec{L}.\vec{S}$ où :

$$H_0 = \frac{\vec{p}^2}{2\mu} + V(r) \ , \quad V(r) = -\frac{e'^2}{r} \ , \quad A(r) = \frac{g_\mathrm{e}}{4\mu^2c^2}\,\frac{1}{r}\,\frac{\mathrm{d}V}{\mathrm{d}r} \ . \qquad (23.228)$$

   (a) Le moment cinétique total $\vec{J} = \vec{L} + \vec{S}$ est une constante du mouvement (le vérifier). Quelles sont les valeurs possibles de j ?

   (b) En déduire au premier ordre la variation d'énergie ΔE_{nj} des états $|njml\,1/2\rangle$ résultant du couplage spin-orbite.

   (c) Trouver la structure fine[68] du niveau 2P ($n = 2, l = 1$). α désignant la constante de structure fine et sachant que :

$$\langle R_{21}|A(r)|R_{21}\rangle = \frac{1}{48}\hbar^{-2}\alpha^4\mu c^2 \ , \qquad (23.229)$$

   calculer en eV et en MHz les déplacements d'énergie.

[67]Un autre exemple : les sous-niveaux hyperfins du fondamentale de l'ion Be^+ (voir chapitre 21, section 21.1, fig. 21.2).

[68]Le calcul effectué ici doit être compris comme un simple exemple d'application de la théorie des perturbations. Le traitement complet de la structure fine doit incorporer d'autres corrections du même ordre de grandeur que celles considérées ici (voir [25], chapitre XII, C.3)

2. L'atome est maintenant soumis de surcroît à un petit champ magnétique $\vec{\mathcal{B}}$ dirigé le long de Oz. En délaissant les termes diamagnétiques, le Hamiltonien est :

$$H = H_0 + A(r)\vec{L}.\vec{S} - \mu_{\mathrm{B}}\hbar^{-1}(L_z + g_{\mathrm{e}}S_z)\mathcal{B} \equiv H_0 + H' \qquad \left(\mu_{\mathrm{B}} = \frac{e\hbar}{2\mu} < 0\right) \ . \quad (23.230)$$

Le couplage Zeeman est traité comme une perturbation sur les niveaux de structure fine obtenus en 1. À l'aide du théorème de Wigner - Eckart, exprimer la variation d'énergie $\Delta E_{\mathrm{Zeeman}}$ du niveau $|n\,j\,m\,l\,1/2\rangle$. Calculer en eV les déplacements Zeeman pour $\mathcal{B} = 10^{-2}$ T ($|\mu_{\mathrm{B}}| \simeq 5{,}8 \times 10^{-5}$ eV/T) par rapport aux énergies en champ nul, $E_{2P_{3/2}}$ et $E_{2P_{1/2}}$.

3. Lorsque le champ magnétique devient plus intense, il n'est plus légitime de négliger le couplage Zeeman entre les deux types d'états $j = l \pm 1/2$; expliquer pourquoi.

   Dans la suite, on considère exclusivement le niveau 2P. Pour simplifier, on note comme suit les états $|2\,j\,m\,l\,s\rangle$ pertinents :

$$|2\,1/2\,m\,1\,1/2\rangle \equiv |1/2\,m\rangle \ , \quad |2\,3/2\,m\,1\,1/2\rangle \equiv |3/2\,m\rangle \ , \quad (23.231)$$

   et on introduit un H' effectif, projection de H' dans le sous-espace 2P :

$$H'_{\mathrm{eff}} = \frac{a}{2}(\vec{J}^2 - \vec{L}^2 - \vec{S}^2) + b(J_z + S_z) \ . \quad (23.232)$$

   (a) Identifier les constantes a et b, caractéristiques du sous-espace considéré.

   (b) Préciser les éléments de matrice de S_z qui sont nuls.

   (c) En déduire les énergies en présence du champ pour $m = \pm 3/2$, $E_{2P_{3/2},m=\pm 3/2}$.

   (d) Écrire le déterminant donnant les valeurs propres pour $m = \pm 1/2$ et en déduire[69] les valeurs propres $E_{\pm,m=\pm 1/2}$.

   (e) Sur un même graphique, porter la variation des énergies en fonction du champ magnétique. Comparer à l'approximation utilisée en 2 (effectuer un développement limité). Expliquer pourquoi, dans l'approximation considérée, seuls les niveaux ayant des m différents ont la possibilité de se croiser.

23.4.12 Atome d'hydrogène dans deux champs croisés

Un atome d'hydrogène dans l'état $n = 2$ est plongé dans deux champs électrique $\vec{\mathcal{E}}$ et magnétique $\vec{\mathcal{B}}$ perpendiculaires, tous deux uniformes et constants. On prend la direction de $\vec{\mathcal{B}}$ comme axe de quantification Oz, et on notera g_{e} le facteur anormal de l'électron.

1. Former le Hamiltonien complet.

[69] Observer qu'un élément *non-diagonal* linéaire en $\mathcal{B}$ produit une correction du *second* ordre (*quadratique* en $\mathcal{B}$).

2. Pourquoi peut-on étudier indépendamment l'un de l'autre les deux sous-espaces $m_s = \pm\frac{1}{2}$? Dans la suite, on ne considère que les états ayant $m_s = +\frac{1}{2}$.

3. Calculer, au premier ordre, les déplacements d'énergie induits par les deux champs dans les trois cas suivants :

   (a) couplages électrique et magnétique du même ordre de grandeur ;

   (b) couplage électrique très petit devant le couplage magnétique ;

   (c) couplage magnétique très petit devant le couplage électrique.

23.4.13 Effet Zeeman sur un atome alcalin

On sait (voir problème 19.3.15) qu'une façon de représenter l'effet des électrons internes sur l'électron périphérique consiste à compléter l'interaction de Coulomb $-Z\frac{e'^2}{r}$ par un terme variant comme $\frac{1}{r^2}$. On pose plus précisément :

$$V(r) = -Z\frac{e'^2}{r} + C\frac{\hbar^2}{2\mu r^2} \quad . \tag{23.233}$$

Quel est le signe de la constante C ?

1. Dans cette partie, on ignore les effets liés au spin ; soit $E_n \stackrel{\text{déf}}{=} -\frac{Z^2\alpha^2}{2n^2}mc^2$ les énergies de l'atome quand $C = 0$.

   (a) Former l'équation radiale pour la fonction $u(r)$.

   (b) En admettant que $u(r) \propto r^\beta + ...$ quand $r \ll a_0$, trouver l'exposant β en fonction de l et de C.

   (c) On cherche la fonction radiale sous la forme d'une série entière en r. En reprenant l'argument standard, trouver les niveaux d'énergie E_{nl}.

   (d) Comparer le résultat précédent en utilisant la théorie des perturbations au premier ordre ; on donne l'élément de matrice :

$$\langle\psi_{nlm}|\frac{1}{r^2}|\psi_{nlm}\rangle = \frac{2}{n^3(2l+1)a_0^2} \quad . \tag{23.234}$$

2. On introduit maintenant l'interaction spin-orbite liée au moment magnétique de l'électron $\vec{\mu} = \gamma_e\vec{S}$, se traduisant par le couplage magnétique $V_{\text{magn}} = A(r)\vec{L}.\vec{S}$ (voir (20.20)).

   (a) Quelles sont les observables commutant avec le Hamiltonien ainsi complété ?

   (b) Quel est l'ordre de grandeur de $V_{\text{magn}}/|E_1|$?

   (c) Calculer, au premier ordre, la correction d'énergie due à V_{magn} pour les niveaux $n = 1$ et $n = 2$. En déduire le schéma des niveaux d'énergie.

(d) Former les états propres communs à H, $\vec{J}^2$, J_z, $\vec{L}^2$ et $\vec{S}^2$.

3. On ajoute maintenant un très petit champ magnétique de module $\mathcal{B}$, autorisant à traiter chaque multiplet indépendamment des autres.

   (a) Calculer les écarts Zeeman.

   (b) Écrire la condition sur $\mathcal{B}$, C et Z pour que l'approximation précédente soit valide.

23.4.14 Effet Zeeman sur un oscillateur harmonique isotrope : traitements perturbatif et exact

On considère un électron (masse m, charge e) lié dans un puits harmonique isotrope dans $\mathbb{R}^3$ (pulsation propre ω_0), soumis de surcroît à un champ magnétique $\vec{\mathcal{B}}$ définissant l'axe Oz. Il s'agit de trouver les déplacements des niveaux d'énergie, soit par un traitement exact [190], soit par perturbation quand le champ est faible. On posera :

$$\omega_{\mathrm{L}} \stackrel{\text{déf}}{=} \frac{|e|\mathcal{B}}{2m} \quad \text{(pulsation de Larmor)} \ , \qquad \omega \stackrel{\text{déf}}{=} \sqrt{\omega_0^2 + \omega_{\mathrm{L}}^2} \ , \tag{23.235}$$

et :

$$k_0 \stackrel{\text{déf}}{=} \sqrt{\frac{m\omega_0}{\hbar}} \ , \qquad k \stackrel{\text{déf}}{=} \sqrt{\frac{m\omega}{\hbar}} \ . \tag{23.236}$$

1. *Traitement exact*

   (a) Écrire l'expression du Hamiltonien H, en y faisant figurer le moment cinétique orbital $\vec{L}$; on pourra noter $r_\perp \stackrel{\text{déf}}{=} x^2 + y^2$.

   (b) Montrer que H se décompose en $H = H_\parallel + H_\perp$, avec $[H_\parallel, H_\perp] = 0$; en déduire que les fonctions propres sont de la forme $\psi(\vec{r}) = \psi_\parallel(z)\psi_\perp(x, y)$, et que l'énergie propre E associée à $\psi(\vec{r})$ s'écrit $E = E_\perp + E_\parallel$.

   (c) On résout maintenant successivement les deux problèmes découplés.

      i. *Mouvement parallèle à Oz*

         A. On définit les opérateurs de création et d'annihilation :

$$a_z \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}\big(k_0 z + \frac{\mathrm{i}}{\hbar k_0} p_z\big) \ , \quad a_z^\dagger \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}\big(k_0 z - \frac{\mathrm{i}}{\hbar k_0} p_z\big) \ . \tag{23.237}$$

         Combien vaut $[a_z, a_z^\dagger]$?

         B. Notant n_z les valeurs propres de l'opérateur $N_z \stackrel{\text{déf}}{=} a_z^\dagger a_z$, en déduire $E_{\parallel,\, n_z}$.

         C. Exprimer $|\psi_{\parallel,\, n_z}\rangle$ en fonction de l'état fondamental $|0\rangle_z$.

ii. *Mouvement dans le plan xOy*

A. On définit de façon analogue ($u = x,\, y$) :

$$a_u \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}\big(ku + \frac{\mathrm{i}}{\hbar k}p_u\big) \quad , \quad a_u^\dagger \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}\big(ku - \frac{\mathrm{i}}{\hbar k}p_u\big) \quad , \tag{23.238}$$

puis les opérateurs $a_\mathrm{r} \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}\big(a_x - \mathrm{i}a_y\big)$, $a_\mathrm{l} \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}\big(a_x + \mathrm{i}a_y\big)$ et leurs hermitiques conjugués $a_\mathrm{r}^\dagger$ et $a_\mathrm{l}^\dagger$. Quelles sont les relations de commutation entre ces opérateurs ?

B. On introduit les opérateurs $N_\mathrm{r} \stackrel{\text{déf}}{=} a_\mathrm{r}^\dagger a_\mathrm{r}$ et $N_\mathrm{l} \stackrel{\text{déf}}{=} a_\mathrm{l}^\dagger a_\mathrm{l}$, et on note n_r, n_l leurs valeurs propres. Montrer que $H_\perp$ s'écrit :

$$H_\perp = \hbar\omega(N_\mathrm{r} + N_\mathrm{l} + 1) + \hbar\omega_\mathrm{L}(N_\mathrm{r} - N_\mathrm{l}) \ , \tag{23.239}$$

C. En déduire l'expression de $E_{\perp,\, n_\mathrm{r} n_\mathrm{l}}$, et celle de $|\psi_{\perp,\, n_\mathrm{r} n_\mathrm{l}}\rangle$ en fonction de $|00\rangle$, fondamental d'un oscillateur harmonique plan.

2. *Traitement par perturbation*

   Résoudre le même problème par perturbation, en examinant les trois premiers niveaux d'énergie de l'oscillateur non-perturbé.

3. *Comparaison des deux traitements*

   (a) Afin de préciser les situations extrêmes (champ faible et champ fort), faire les développements limités pertinents des énergies exactes.

   (b) En déduire aussi précisément que possible la variation en fonction du champ des énergies exactes des premiers niveaux, et observer les croisements de niveaux persistents.

   (c) Comparer les développements à petit champ avec le traitement perturbatif.

23.4.15 Matrice densité de Bloch

$\{\psi_n(\vec{r})\}_n$ étant l'ensemble des fonctions propres d'un Hamiltonien H à une particule, de valeurs propres E_n, on définit la matrice densité de Bloch par l'égalité :

$$C(\vec{r}\,',\, \vec{r},\, \beta) \stackrel{\text{déf}}{=} \sum_n \psi_n^*(\vec{r}\,')\psi_n(\vec{r})\,\mathrm{e}^{-\beta E_n} \quad , \quad \beta = (k_\mathrm{B}T)^{-1} \tag{23.240}$$

1. Montrer que C satisfait l'équation dite *de Bloch* :

$$\frac{\partial C}{\partial \beta} = -HC \ . \tag{23.241}$$

   Quelle comparaison pouvez-vous faire avec l'équation de Schrödinger ? Écrire la relation précise entre C et le propagateur U.

2. Combien vaut $C(\vec{r}\,', \vec{r}, 0)$?

3. Trouver la solution $C_0(\vec{r}\,', \vec{r}, \beta)$ de l'équation de Bloch pour une assemblée de fermions libres, chacun d'entre eux ayant le Hamiltonien $H_0 \stackrel{\text{déf}}{=} \frac{\vec{p}^{\,2}}{2m}$. Faire le lien avec le propagateur d'une particule libre (voir Tome I, éq. (4.129)).

4. Pour des fermions sans interaction mutuelle mais soumis à un champ de force, le Hamiltonien à une particule est $H = H_0 + V(\vec{r})$. L'équation de Bloch s'écrit alors :

$$\frac{\partial C}{\partial \beta} + H_0 C = -VC \ , \tag{23.242}$$

5. Montrer que la solution de (23.242) peut s'écrire :

$$C(\vec{r}\,', \vec{r}, \beta) = C_0(\vec{r}\,', \vec{r}, \beta) - \int \mathrm{d}^3 r_1 \int_0^\beta \mathrm{d}\beta' \, C_0(\vec{r}\,', \vec{r}_1, \beta - \beta') V(\vec{r}_1) C(\vec{r}_1, \vec{r}, \beta') \ . \tag{23.243}$$

6. À l'ordre le plus bas, on a donc :

$$C(\vec{r}\,', \vec{r}, \beta) \simeq C_0(\vec{r}\,', \vec{r}, \beta) - \int \mathrm{d}^3 r_1 \int_0^\beta \mathrm{d}\beta' \, C_0(\vec{r}\,', \vec{r}_1, \beta - \beta') V(\vec{r}_1) C_0(\vec{r}_1, \vec{r}, \beta') \ . \tag{23.244}$$

Soit $\Gamma(\vec{r}\,', \vec{r}, z)$ la transformée de Laplace de $C(\vec{r}\,', \vec{r}, \beta)$, et $\Gamma_0(\vec{r}\,', \vec{r}, z)$ celle de $C_0(\vec{r}\,', \vec{r}, \beta)$:

$$\Gamma(\vec{r}\,', \vec{r}, z) \stackrel{\text{déf}}{=} \int_0^{+\infty} C(\vec{r}\,', \vec{r}, \beta) \, \mathrm{e}^{-z\beta} \, \mathrm{d}\beta \ . \tag{23.245}$$

Trouver[70] $\Gamma_0(\vec{r}\,', \vec{r}, z)$.

7. Utiliser le théorème de convolution pour écrire l'expression intégrale de $\Gamma(\vec{r}\,', \vec{r}, z)$ à cet ordre d'approximation.

8. En déduire $\Gamma(\vec{r}\,', \vec{r}, z)$ pour un potentiel de contact $V(\vec{r}) = g\delta(\vec{r})$.

9. Pour ce même potentiel, trouver l'expression exacte de $\Gamma(\vec{r}\,', \vec{r}, z)$ et comparer au résultat précédent.

10. Sachant que l'inversion de Laplace s'effectue en faisant la chasse aux singularités, quelle différence qualitative majeure existe entre le résultat approché et le résultat exact ? Discuter très qualitativement les deux cas $g > 0$ et $g < 0$.

[70] On donne l'intégrale $\int_0^{+\infty} x^{-3/2} \, \mathrm{e}^{-ax+b/x} \, \mathrm{d}x = \sqrt{\frac{\pi}{b}} \, \mathrm{e}^{-2\sqrt{ab}}$.

physique

Chapitre 24

Théorie des perturbations dépendant du temps

Le but de ce chapitre est d'exposer les bases de la théorie des perturbations dépendant du temps et de présenter ses premières applications à quelques cas simples mais importants. (résonances, Règle d'or de Fermi, règles de sélection, effet photoélectrique). Le chapitre se termine sur la démonstration du théorème fondamental de Gell-Mann et Low.

24.1 Présentation générale

L'étude expérimentale d'un système, noté (Sy) dans la suite, consiste presque toujours à faire interagir ce système (que l'on peut appeler *cible*) avec une sonde extérieure dont les caractéristiques physiques sont contrôlées. Très souvent, la sonde est un faisceau de particules d'énergie définie (excitation *monochromatique*) ; si ces dernières sont douées d'un caractère vectoriel intrinsèque (un spin, par exemple), le faisceau peut, si besoin est, être préparé dans un état de polarisation bien déterminée (lumière polarisée pour sélectionner certaines transitions atomiques, neutrons polarisés pour mettre en évidence un ordre magnétique en phase condensée, etc.). Le faisceau diffusé[1], en raison de son interaction avec la cible, présente des altérations dont l'analyse permet précisément de remonter aux propriétés intrinsèques de la cible ; en quelque sorte, le faisceau diffusé constitue une image au sens large de la cible.

[1]En français, le vocable diffusion a un double sens :

- diffusion au sens d'éparpillement, c'est le sens employé premièrement en Théorie des collisions ;
- diffusion au sens de l'équation de la diffusion déduite de la loi de Fick ($\partial P/\partial t = D\Delta P$).

En anglais, on dispose de deux substantifs, respectivement *scattering* et *diffusion*.

Dans la situation la plus courante, le faisceau est constitué soit de particules matérielles, soit de photons et l'expérience consiste à mesurer des sections efficaces. La dépendance de ces dernières vis-à-vis de l'énergie (ou de la fréquence) et de l'angle de diffusion fournit de précieux renseignements sur l'interaction entre les projectiles et la cible et permet de ce fait de mieux connaître la dynamique interne de celle-ci – en définitive d'affiner la description théorique (fondamentale ou phénoménologique) des interactions élémentaires, ou des interactions effectives lorsque les projectiles *moyennent* d'une certaine façon une partie des degrés de liberté de la cible.

Traditionnellement et le plus souvent, on parle de *collisions* quand les projectiles sont des particules matérielles, et de *spectroscopie* (nucléaire, atomique, moléculaire, du solide, etc.) quand ce sont des photons qui sont utilisés. Cette terminologie est quelque peu arbitraire : on sait bien que toutes les particules ont un comportement complexe qui ne permet pas une distinction instinctive aussi nette, héritée des conceptions anciennes. Suivant les circonstances précises d'une expérience, une même particule (l'électron par exemple) se comportera *plutôt* comme une onde *ou* comme une petite bille – ce sont des cas extrêmes – mais plus souvent les deux aspects ondulatoire et corpusculaire se superposent, ou plutôt s'imbriquent l'un dans l'autre.

Les applications de la spectroscopie sont innombrables et en font un outil d'usage universel pour l'étude de la matière, qu'il s'agisse de recherche fondamentale ou appliquée. Les mesures utilisant la spectroscopie atomique sont d'une précision diabolique : c'est grâce à certaines qu'il a été possible de vérifier expérimentalement avec une précision rarement atteinte les prévisions de l'Électrodynamique quantique, notamment l'écart $g-2$ du facteur anormal de l'électron. Dans un tout autre champ d'application, les méthodes d'imagerie – telles qu'on les utilise aujourd'hui en Médecine par exemple – sont fondées sur l'aptitude de la matière à réagir sélectivement et de façon spécifique à une petite perturbation dûment contrôlée.

Le problème théorique peut très souvent être formulé comme suit. Le système (Sy), caractérisé par son Hamiltonien H_0, est à un certain instant dans l'un de ses états propres $|\psi_\mathrm{i}\rangle$, constituant un état stationnaire[2]. À partir d'un instant t_0, (Sy) est soumis à une sonde *extérieure* non quantifiée ; le couplage entre (Sy) et cette perturbation est représentée par un opérateur $V(t)$ dépendant du temps où apparaissent conjointement une (ou plusieurs) observables de (Sy) et les grandeurs physiques caractéristiques de la sonde. Pour $t \geq t_0$, l'état du système (Sy) obéit à l'équation de Schrödinger :

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = [H_0 + V(t)]\,|\Psi(t)\rangle \ . \tag{24.1}$$

Comme, d'une façon générale, l'état initial $|\psi_\mathrm{i}\rangle$ n'est pas propre[3] de $H(t)$, cet état, qui serait stationnaire en l'absence de perturbation, va évoluer sous l'effet de $V(t)$. Il est toujours loisible de garder comme base de l'espace des états de (Sy), $\mathcal{E}$, les états propres de H_0, les $\{|E_k\rangle\}_k$; dès lors, l'effet de la perturbation est d'induire des transitions entre

[2]Cette hypothèse n'est pas nécessaire ; elle est admise pour fixer les idées et aussi parce qu'elle correspond à la situation la plus courante.

[3]Si c'était le cas, le problème serait sans grand intérêt.

l'état initial $|\psi_\mathrm{i}\rangle$ et les autres états propres de (Sy). On peut se représenter l'effet de la perturbation comme une promenade forcée du système dans son espace d'états $\mathcal{E}$.

La question typique que l'on se pose alors est la suivante : $|\psi_\mathrm{f}\rangle$ désignant l'un quelconque des états propres $\{|E_k\rangle\}_k$, quelle est la probabilité $P_{\mathrm{i}\rightarrow\mathrm{f}}(t)$ pour que le système se retrouve à l'instant t dans l'état $|\psi_\mathrm{f}\rangle$ sous l'effet de la perturbation ? Cette question a une réponse et une seule : l'équation (24.1) est du premier ordre en temps et on s'est donné un état initial, $|\psi_\mathrm{i}\rangle$. En supposant effectuée l'intégration en temps, donc ayant déterminé le vecteur d'état à l'instant t, $|\Psi(t)\rangle$, la probabilité cherchée s'exprime, en vertu des postulats, sous la forme du module au carré d'une certaine amplitude de probabilité :

$$P_{\mathrm{i}\rightarrow\mathrm{f}}(t) = |A_{\mathrm{i}\rightarrow\mathrm{f}}(t)|^2 \equiv |\langle\psi_\mathrm{f}|\Psi(t)\rangle|^2 \qquad (|\Psi(t=t_0)\rangle = |\psi_\mathrm{i}\rangle) \ . \tag{24.2}$$

Si $U(t, t_0)$ désigne l'opérateur d'évolution associé à $H(t)$ entre les instants t_0 et t (voir Tome I, chapitre 14, éq. (14.17)), l'amplitude de probabilité $A_{\mathrm{i}\rightarrow\mathrm{f}}$ est donnée par :

$$\boxed{A_{\mathrm{i}\rightarrow\mathrm{f}}(t) = \langle\psi_\mathrm{f}|U(t,t_0)|\psi_\mathrm{i}\rangle} \tag{24.3}$$

Les amplitudes étant trouvées, on peut alors en déduire notamment la variation d'énergie (moyenne) du système résultant du passage d'une perturbation provisoire ; en désignant par $A_{\mathrm{i}\rightarrow k}(t)$ l'amplitude de probabilité de transition entre l'état initial d'énergie E_i et l'état propre $|E_k\rangle$ d'énergie E_k, la variation d'énergie est :

$$\Delta E \equiv E_\mathrm{finale} - E_\mathrm{i} = \Big(\sum_k E_k\,|A_{\mathrm{i}\rightarrow k}(t)|^2\Big) - E_\mathrm{i} = \sum_k\,(E_k - E_\mathrm{i})\,P_{\mathrm{i}\rightarrow k} \ . \tag{24.4}$$

La dernière égalité résulte du fait que la somme des probabilités de transitions de l'état initial aux différents états finals doit être égale à 1, $\sum_k P_{\mathrm{i}\rightarrow k} = 1$, puisque le système doit bien se trouver quelque part.

Le problème est ainsi bien posé, mais n'est malheureusement pas soluble exactement, mis à part quelques cas simples – qui demeurent cependant très instructifs. Comme exemples de problèmes exactement solubles, on peut citer : un oscillateur harmonique chargé soumis à un champ électrique variable de façon quelconque dans le temps (Tome I, problème 16.6.7, et aussi problème 24.7.9), un spin 1/2 soumis à une perturbation harmonique[4] (problème 18.7.11), deux spins $S_i = 1/2$ en interaction par un couplage du genre $a(t)\vec{S}_1.\vec{S}_2$ où $a(t)$ est une fonction réelle quelconque (problème 24.7.2).

Fort heureusement, et c'est presque une nécessité méthodologique, la perturbation qui permet de sonder le système d'intérêt doit rester *petite*, en un sens à préciser. En effet, c'est le système lui-même que l'on veut observer et il convient donc de le déranger le moins possible. Quand il en a été décidé ainsi, il y a forcément dans le terme de couplage $V(t)$ un petit paramètre sans dimension, g, et on peut espérer pouvoir construire un développement de l'amplitude $A_{\mathrm{i}\rightarrow\mathrm{f}}(t)$ en série entière de g ; si $g \ll 1$, le ou les tout

[4] à condition de délaisser les terme antirésonnants.

premiers termes seront en général suffisants. Ceci est l'esprit de toutes les méthodes de perturbation, classiques ou quantiques d'ailleurs[5].

La décomposition présupposée du Hamiltonien total $H(t) = H_0 + V(t)$ est toujours le résultat d'une approximation. Fondamentalement, on suppose qu'il est possible de séparer et d'identifier clairement d'une part le système (Sy), décrit par H_0, et la perturbation extérieure, représentée par $V(t)$. Il y a bel et bien une dissymétrie entre l'objet d'observation et l'outil utilisé pour cette observation. De fait, la perturbation est considérée comme un champ extérieur donné qui bouscule le système mais n'en subit aucune contre-réaction ; on se doute physiquement que ceci sera licite si le champ extérieur est un grand système, constituant notamment un réservoir infini d'énergie, totalement désincarné physiquement et supposé complètement décrit par un (petit) nombre de grandeurs physiques données une fois pour toutes. En d'autres termes, seul le système (Sy) fait l'objet d'une description quantique, le champ étant décrit classiquement ; c'est pourquoi une telle approche est appelée *semi-classique*.

Construire une justification théorique de ce type d'approche – tout en restant à un niveau élémentaire[6] – c'est énoncer les conditions permettant de se contenter de l'approche semi-classique et, *a contrario*, préciser les circonstances où une description complètement quantique est requise. Dans ce but, on se met d'emblée dans le cadre le plus rigoureux, en quantifiant à la fois le système d'intérêt et la perturbation avec sa source, appelée champ dans la suite ; le Hamiltonien du supersystème $[(Sy) + \text{champ}]$ – qui constitue alors un système isolé – peut toujours être mis sous la forme :

$$H = H_0 + H_{\text{champ}} + H_{\text{int}} \quad . \tag{24.5}$$

H, décrivant un système isolé, est alors indépendant du temps et est fonction des degrés de liberté de (Sy), notés collectivement $\{x\}$, et de ceux du champ, $\{X\}$. H_0 ne dépend que de $\{x\}$, H_{champ} ne dépend que de $\{X\}$; quant à H_{int}, s'agissant de l'opérateur décrivant l'interaction entre les deux sous-systèmes, c'est bien sûr une fonction des deux jeux de variables $\{x\}$ et $\{X\}$.

Le système (Sy) a son propre espace d'états, $\mathcal{E}_{Sy}$, celui du champ est $\mathcal{E}_{\text{champ}}$; le système total a pour espace d'états le produit tensoriel $\mathcal{E}_{Sy} \otimes \mathcal{E}_{\text{champ}}$. H_{int} couple entre eux ces états produits tensoriels – qui ne sont propres, donc stationnaires, qu'en l'absence précisément de H_{int}, quand chacun des partenaires évolue sans se soucier de l'autre, faute d'interaction.

La dynamique du système total, un état initial étant donné, est déterminée par les équations de Heisenberg pour les variables $\{x\}$ et $\{X\}$; formellement :

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,x_{\mathrm{H}}(t) = [x, H_0]_{\mathrm{H}}+[x, H_{\text{int}}]_{\mathrm{H}} \ , \qquad \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,X_{\mathrm{H}}(t) = [X, H_{\text{champ}}]_{\mathrm{H}}+[X, H_{\text{int}}]_{\mathrm{H}} \ . \tag{24.6}$$

[5]Bien sûr, il existe des cas où la perturbation est délibérément choisie intense, par exemple pour observer les fragments d'un système qui n'aura pas résisté et se sera décomposé ; cette situation ne relève en aucune façon d'une théorie de perturbation, au sens précisé ci-dessus.

[6]La justification ci-dessous est élémentaire au sens où elle ignore les cas où H_0 contient déjà en fait certaines modifications prenant la perturbation en compte (*renormalisation* de la masse, de la charge, ...).

Le second membre de chaque équation contient d'une part l'évolution libre (premier terme), d'autre part un terme de couplage impliquant les opérateurs des deux partenaires, (Sy) et champ ; par exemple, l'équation pour x est plus précisément de la forme :

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\, x_{\mathrm{H}}(t) = [x, H_0]_{\mathrm{H}} + W(\{x_{\mathrm{H}}\}, \{X_{\mathrm{H}}\}) \tag{24.7}$$

où W est un opérateur agissant dans l'espace produit tensoriel $\mathcal{E}$ du supersystème. La description semi-classique sera justifiable si, pour une raison à préciser, les variables $\{X_{\mathrm{H}}\}$ peuvent être remplacées par leurs valeurs moyennes. Ceci implique que les *fluctuations* quantiques du champ peuvent être négligées ; alors, on obtient :

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\, x_{\mathrm{H}}(t) = [x, H_0]_{\mathrm{H}} + W(\{x_{\mathrm{H}}\}, \{\langle X_{\mathrm{H}}\rangle\}) \equiv [x, H_0]_{\mathrm{H}} + F(\{x_{\mathrm{H}}\}; t) \tag{24.8}$$

où maintenant la représentation de Heisenberg, définie par référence au Hamiltonien total[7] H, peut en réalité être élaguée de tous les termes propres au champ. Cependant, la théorie reste ici purement mécanique[8], et il sera toujours possible de récrire le terme d'interaction simplifié F sous la forme d'un commutateur :

$$F(\{x_{\mathrm{H}}\}; t) \equiv [x, V(t)]_{\mathrm{H}} \ . \tag{24.9}$$

Dans ces conditions, l'équation (24.6) pour x devient :

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\, x_{\mathrm{H}}(t) = [x, H_0 + V(t)]_{\mathrm{H}} \ ; \tag{24.10}$$

il est naturel de définir un Hamiltonien dépendant du temps, $H(t) = H_0 + V(t)$, décrivant l'interaction du système en interaction avec le champ extérieur décrit classiquement.

Donnons quelques exemples. Soit un système quantique bien identifié (atome, noyau, molécule), couplé à un champ électromagnétique piloté de l'extérieur. Dans la situation la plus courante, l'atome est initialement dans son état fondamental et, sous l'effet du champ, peut grimper dans ses états excités ; c'est l'image la plus simple expliquant l'existence de raies spectrales d'absorption, caractérisées par les fréquences déduites de la loi de Bohr :

$$\nu_{nm} = \frac{1}{h}\,(E_n - E_m) \tag{24.11}$$

et par des intensités plus ou moins grandes que l'on essaie de déduire des caractéristiques quantiques de l'atome.

Quand on se souvient que le champ électromagnétique est constitué de photons, le procédé semi-classique peut sembler très cavalier. Une justification un peu simpliste

[7] Si tel n'était pas le cas, il y aurait encore des opérateurs du champ dans les équations simplifiées. En réalité, à partir du moment où les opérateurs $\{\langle X_{\mathrm{H}}\rangle\}$ ont été remplacés par leurs valeurs moyennes, toutes les grandeurs relatives au champ sont des scalaires ; dans la transformation de Heisenberg définie par H, tous les termes propres au champ disparaissent de fait en faisant commuter les scalaires. Une telle représentation n'est pas la *représentation-interaction* définie plus bas, puisque outre H_0, il subsiste le terme $H_{\mathrm{int}} + H_{\mathrm{champ}}$ moyenné sur tous les degrés de liberté X du champ.

[8] par opposition à une théorie dissipative donnant lieu à une évolution irréversible ; une telle évolution résulte de termes non-hamiltoniens donnant au second membre de l'équation de Heisenberg des termes d'évolution n'ayant pas la forme d'un commutateur.

s'énonce ainsi : avec les sources ordinaires, il y a en général un nombre énorme de photons présents[9] et les fluctuations quantiques sont souvent négligeables[10]. Inversement, il est essentiel de quantifier le champ pour rendre compte d'un phénomène aussi important que l'émission spontanée ; une fois l'atome passé à l'état excité par l'action d'une perturbation transitoire, il n'y reste pas : à l'instar d'un noyau radioactif, il retombe à l'état fondamental au bout d'un temps aléatoire, dont la moyenne (statistique) constitue par définition la durée de vie de l'état excité. La description de ce phénomène échappe totalement à l'approche semi-classique – et pour cause : il se produit en l'absence de champ "extérieur", une situation où le nombre de photons "présents" est *en moyenne* nul. C'est évidemment une situation où les fluctuations vont jouer un rôle éminent. De fait, l'émission spontanée résulte de ce que l'on appelle les "fluctuations du vide", preuve que le champ électromagnétique est bel et bien un objet relevant d'une description quantique ; les modules des champs électrique et magnétique peuvent être nuls, mais la moyenne de leur carré ne peut évidemment pas l'être, d'où l'existence de fluctuations au sens statistique. Le photon peut s'échapper de l'atome excité grâce à une telle fluctuation, un peu à la manière d'une particule classique qui, sous l'effet d'une fluctuation thermique, peut franchir une barrière d'énergie. D'ailleurs, l'argument vaut tout autant si le photon "émis" ne quitte pas le système, ce dernier le réabsorbant : ce jeu permanent d'émission et d'absorption virtuelles est responsable des corrections radiatives des niveaux atomiques (Lamb *shift*), de l'interaction de van der Waals entre deux objets neutres (atomes ou molécules) à grande distance (voir chapitre 25, section 25.4), etc.

Un second exemple, fourni par l'excitation Coulombienne, montre d'ailleurs que le nombre de particules en jeu n'est pas forcément la variable à considérer quand il s'agit de *déquantifier* certains degrés de liberté. Ce phénomène est la transition vers un état excité d'un atome ou d'une molécule, sous l'effet de l'interaction de Coulomb, lors du passage à proximité[11] d'une particule de charge Q. La description la plus simple consiste à ne retenir que l'interaction entre les particules chargées q_i de l'atome avec le champ électrostatique créé par la particule chargée passant au voisinage ; on écrit ainsi :

$$V(t) = \sum_i \frac{q_i Q}{4\pi\varepsilon_0 \|\vec{R}(t) - \vec{r}_i\|} \tag{24.12}$$

où $\vec{R}$ est la coordonnée du projectile, considéré comme un point matériel classique, $\{\vec{r}_i\}$

[9] Une lampe de 100 W émettant dans le jaune donne, à 1 m de distance, un courant de photons de l'ordre de 2×10^{15} par seconde et par cm^2.

[10] Cette dernière affirmation est usuelle dans un contexte purement classique : quand un grand nombre de variables aléatoires entre en jeu, les fluctuations régressent et, à la limite, les processus deviennent certains. On peut aussi invoquer ici le fait que lorsque le champ contient beaucoup de photons, il se trouve de fait dans la limite des grands nombres quantiques.

Il ne faut toutefois pas en déduire que l'intervention d'un grand nombre de particules justifie toujours d'oublier les effets quantiques. Il existe des transitions de phase (exemples : la condensation de Bose, la supraconductivité) dont l'existence repose sur la symétrie spécifique des fonctions d'onde selon qu'il s'agit de bosons ou de fermions ; d'une part cette symétrie fondamentale fonctionne par tout ou rien et ne peut être ignorée, d'autre part, une transition de phase ne peut survenir, *stricto sensu*, qu'après avoir pris la limite thermodynamique : toute fonction de partition est une certaine somme de fonctions analytiques ; de ce fait elle ne peut présenter de singularités que si cette somme est en fait une *série*.

[11] Ce phénomène d'excitation des atomes est le principal mécanisme par lequel une particule chargée de haute énergie (quelques MeV) perd peu à peu son énergie en traversant la matière ordinaire (voir Tome I, chapitre 2, particulièrement problème 2.3.5).

celles des électrons atomiques. Cette description se doit d'être justifiée ; notamment il faut que, pour la cible, la structure physique du projectile soit invisible. Pour qu'il en soit ainsi, plusieurs conditions doivent être satisfaites. Par exemple, si la particule chargée est elle-même un électron, il y a visiblement un problème vis-à-vis de l'indiscernabilité entre cet électron et ceux de l'atome – sauf bien sûr si celui qui constitue le projectile passe effectivement très loin, auquel cas il peut de fait être distingué de ceux de la cible (non-recouvrement des fonctions d'onde). D'un autre côté, on peut se poser la question de l'effet inverse : l'excitation Coulombienne peut tout autant s'effectuer dans l'autre sens, la cible excitant le projectile supposé muni d'une structure interne[12]. Ceci ne peut se produire si les niveaux internes du projectile sont très bien séparés en énergie, caractérisés par une échelle très grande devant les énergies internes de la cible et devant la perturbation. Au total, la description ci-dessus doit bien fonctionner pour une particule assez massive (pour avoir une petite longueur d'onde associée) et très robuste (par exemple, une particule α traversant un gaz).

♦ *Remarque*

Dans le cas d'une description semi-classique où $H(t) = H_0 + V(t)$, il est loin d'être évident que $H(t)$ représente l'énergie du système. En effet, il s'agit toujours, comme on l'a vu, d'une réduction d'un problème indépendant du temps (où, pour de bonnes raisons, un autre système a été désincarné physiquement). Sur le plan des principes, il existe donc plus fondamentalement une interaction notée H_{int} dans l'analyse ci-dessus. À partir de ce fait, il n'est plus possible à proprement parler de parler de l'énergie de l'une des parties du système, sauf à convenir que H_{int} doit être coupé en deux moitiés égales dans un jugement de Salomon que rien ne justifie à ce niveau.

D'un autre côté, dans la formulation semi-classique, il est parfaitement possible de définir sans aucune ambiguïté la variation d'énergie d'un système donné sous l'effet du passage d'une perturbation de durée finie. Si t_{avant} et $t_{\text{après}}$ désignent deux instants où la perturbation n'est pas présente et encadrant son passage, on a :

$$E(t_{\text{avant}}) - E(t_{\text{après}}) \equiv E_{\text{f}} - E_{\text{i}} = \Delta E \tag{24.13}$$

où la variation d'énergie ΔE est donnée par l'expression (24.4). En toutes circonstances, comme d'habitude, ce sont les *différences* d'énergie qui comptent. ♦

24.2 Calcul systématique de l'opérateur d'évolution et approximations successives

Soit un système quantique de Hamiltonien H_0, d'états propres $\{|E_k\rangle\}_k$, soumis à partir d'un instant pris comme origine des temps à une perturbation extérieure se couplant au

[12] Cette question ne se pose évidemment pas si le projectile est élémentaire au sens strict, comme l'électron.

système par l'opérateur $gV(t)$; le nombre g est utile pour compter les puissances le cas échéant – physiquement c'est une constante de couplage sans dimension (la constante de structure fine par exemple), conventionnellement supposée positive, et que l'on peut formellement prendre égale à 1 pour se recaler sur des notations plus usuelles. À $t \geq 0$, le Hamiltonien décrivant le système perturbé est :

$$H(t) = H_0 + gV(t) \ . \tag{24.14}$$

L'état à l'instant t, $|\Psi(t)\rangle$, se déduit de l'opérateur d'évolution $U(t)$, pourvu que l'on connaisse une condition initiale, spécifiée selon $|\Psi(t=0)\rangle = |\psi_{\mathrm{i}}\rangle$ et donnant :

$$\boxed{|\Psi(t)\rangle = U(t)|\psi_{\mathrm{i}}\rangle} \tag{24.15}$$

Résoudre le problème de l'évolution en présence de la perturbation, c'est donc trouver l'opérateur $U(t)$, qui obéit à l'équation :

$$\boxed{\mathrm{i}\hbar \frac{\partial}{\partial t} U(t) = [H_0 + gV(t)]\, U(t)} \tag{24.16}$$

avec la condition initiale $U(t=0) = \mathbf{1}$. C'est cette équation que, dans l'immense majorité des cas, on résout faute de mieux par approximations successives. Dans ce but, il est naturel et commode de mettre en évidence l'évolution libre[13] en posant :

$$U_0(t) \stackrel{\text{déf}}{=} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar} H_0 t} \ , \qquad U_1(t) \stackrel{\text{déf}}{=} U_0^\dagger(t)\, U(t) \quad \Longleftrightarrow \quad U(t) = U_0(t)\, U_1(t) \ . \tag{24.17}$$

En reportant cette forme dans (24.16), on obtient :

$$\mathrm{i}\hbar \left(\frac{\partial U_0}{\partial t} U_1 + U_0 \frac{\partial U_1}{\partial t} \right) = [H_0 + gV(t)]\, U_0 U_1 \tag{24.18}$$

et comme $\mathrm{i}\hbar\partial_t U_0 = H_0 U_0$, il reste :

$$\mathrm{i}\hbar\, U_0(t) \frac{\partial}{\partial t} U_1(t) = gV(t)\, U_0(t) U_1(t) \ ; \tag{24.19}$$

multipliant à gauche par $U_0^\dagger(t)$, on voit que $U_1(t)$ satisfait l'équation :

$$\boxed{\mathrm{i}\hbar \frac{\partial}{\partial t} U_1(t) = gU_0^\dagger(t)\, V(t)\, U_0(t)\, U_1(t) \equiv gV_{\mathrm{I}}(t)\, U_1(t)} \tag{24.20}$$

$V_{\mathrm{I}}(t)$ est appelé *représentation-interaction* de $V(t)$:

$$\boxed{V_{\mathrm{I}}(t) \stackrel{\text{déf}}{=} U_0^\dagger(t)\, V(t)\, U_0(t)} \tag{24.21}$$

tout comme on dit parfois que U_1 est la représentation-interaction de l'opérateur d'évolution, même s'il n'a pas la même forme que V_{I}. Par la définition même de U_1, la condition initiale complétant (24.20) est évidemment :

$$U_1(t=0) = \mathbf{1} \ , \tag{24.22}$$

[13] c'est-à-dire en l'absence de perturbation.

et $V_\text{I}(t)$ satisfait l'équation de Heisenberg :

$$\boxed{\mathrm{i}\hbar\,\frac{\partial V_\text{I}}{\partial t} = [V_\text{I},\, H_0]} \tag{24.23}$$

L'intégration formelle membre à membre de (24.20), compte tenu de la condition initiale (24.22) pour U_1, fournit :

$$U_1(t) = \mathbf{1} + \frac{g}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t_1 V_\text{I}(t_1) U_1(t_1) \ . \tag{24.24}$$

On peut maintenant itérer et écrire :

$$U_1(t) = \mathbf{1} + \frac{g}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t_1 V_\text{I}(t_1)\left[\mathbf{1} + \frac{g}{\mathrm{i}\hbar}\int_0^{t_1} \mathrm{d}t_2 V_\text{I}(t_2) U_1(t_2)\right] \ , \tag{24.25}$$

soit :

$$U_1(t) = \mathbf{1} + \frac{g}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t_1 V_\text{I}(t_1) + \frac{g^2}{(\mathrm{i}\hbar)^2}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\, V_\text{I}(t_1) V_\text{I}(t_2) U_1(t_2) \ . \tag{24.26}$$

En remplaçant à nouveau $U_1(t_2)$ par son expression intégrale (24.24) et en répétant la substitution, on met finalement en évidence le développement infini :

$$\boxed{U_1(t) = \mathbf{1} + \sum_{n=1}^{+\infty} U_1^{(n)}(t)} \tag{24.27}$$

avec :

$$U_1^{(n)}(t) = \left(\frac{g}{\mathrm{i}\hbar}\right)^n \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \ldots \int_0^{t_{n-1}} \mathrm{d}t_n\, V_\text{I}(t_1) V_\text{I}(t_2) \ldots V_\text{I}(t_n) \ . \tag{24.28}$$

Il est clair que chacun des $U_1^{(n)}$ représente un ordre donné de la série de perturbation, $U_1^{(n)}$ étant d'ordre g^n. Il est par ailleurs essentiel de se souvenir que les $V_\text{I}(t_k)$ sont des *opérateurs* et que leur ordre dans le produit ne doit pas être modifié. C'est la raison pour laquelle, en dépit de permutations de variables muettes toujours possibles[14], l'opérateur d'évolution total associé à $H(t)$, $U(t)$, n'est pas égal à $\exp[(\mathrm{i}\hbar)^{-1}\int_0^t \mathrm{d}t'\, H(t')]$.

Il est toutefois possible d'écrire ce développement en série de façon plus symétrique, au prix de quelques complications formelles qui se révéleront utiles (notamment dans la section 24.6). Remarquons d'abord que l'on peut remplacer toutes les bornes supérieures des intégrales par la borne commune t en introduisant les bonnes fonctions échelon-unité, $\theta(t) = 1$ si $t > 0$, $\theta(t) = 0$ si $t < 0$ (à ce stade, il n'est pas nécessaire

[14]Voir le cas $n = 2$, explicité dans l'éq. (14.33) du Tome I.

de définir la valeur de θ en $t = 0$, puisque tout apparaît sous une intégrale) ; l'intégrale multiple de (24.29) peut ainsi s'écrire :

$$\int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \ldots \int_0^t \mathrm{d}t_n \, V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)\ldots V_\mathrm{I}(t_n)\theta(t_1-t_2)\theta(t_2-t_3)\ldots\theta(t_{n-1}-t_n) \ . \quad (24.29)$$

Ceci fait, les intervalles d'intégration étant maintenant les mêmes pour toutes les variables, on peut effectuer n'importe quelle permutation de l'ensemble des n variables muettes d'intégration sans changer la valeur de l'intégrale, ce qui donne $n!$ intégrales toutes égales entre elles ; si P_λ est une permutation quelconque de $\mathcal{S}_n$, l'intégrale dans (24.28) est ainsi égale à :

$$\frac{1}{n!}\int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \ldots \int_0^t \mathrm{d}t_n \sum_{\lambda=1}^{n!} P_\lambda[V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)\ldots V_\mathrm{I}(t_n)\theta(t_1-t_2)\theta(t_2-t_3)\ldots\theta(t_{n-1}-t_n)] \ . \quad (24.30)$$

On voit que la somme des termes permutés est la somme de tous les produits réordonnés chronologiquement suivant $t_1 > t_2 > \ldots > t_n$ de gauche à droite, d'où la définition naturelle de l'opérateur chronologique de Dyson :

$$\mathcal{T}\left[V_\mathrm{I}(t_{i_1})V_\mathrm{I}(t_{i_2})\ldots V_\mathrm{I}(t_{i_n})\right] \stackrel{\text{déf}}{=} V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)\ldots V_\mathrm{I}(t_n) \ \text{si} \ t_1 > t_2 > \ldots > t_n \ . \quad (24.31)$$

Il s'agit d'un opérateur *courtois* qui, agissant sur un produit où les temps sont dans un ordre quelconque, rétablit la chronologie – et les préséances – en faisant agir d'abord les opérateurs les plus anciens. Par exemple, pour $n = 2$:

$$\mathcal{T}\left[V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)\right] \stackrel{\text{déf}}{=} \begin{cases} V_\mathrm{I}(t_1)V_\mathrm{I}(t_2) & \text{si } t_1 > t_2 \\ V_\mathrm{I}(t_2)V_\mathrm{I}(t_1) & \text{si } t_2 > t_1 \end{cases} \ . \quad (24.32)$$

Pour $t_1 = t_2$, il n'y a pas d'ambiguïté puisque $V_\mathrm{I}(t)$ commute évidemment avec lui-même. En introduisant la bonne fonction échelon-unité $\Theta(t)$, on a :

$$\mathcal{T}\left[V_\mathrm{I}(t_{i_1})V_\mathrm{I}(t_{i_2})\right] = V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)\Theta(t_1-t_2) + V_\mathrm{I}(t_2)V_\mathrm{I}(t_1)\Theta(t_2-t_1) \ ; \quad (24.33)$$

il est ici nécessaire de définir la fonction échelon en $t = 0$, puisque l'égalité est écrite indépendamment d'une intégrale ; l'ordre du produit étant sans importance si $t_1 = t_2$, il vient $2\Theta(0) = 1$, soit $\Theta(0) = 1/2$ pour retomber sur ses pieds. L'écriture (24.33) est bien une somme de termes où les indices sont permutés, chaque terme étant multiplié par un produit de $n-1$ fonctions de Heaviside imposant l'ordre chronologique ; en règle générale, on a donc :

$$\mathcal{T}\left[V_\mathrm{I}(t_{i_1})V_\mathrm{I}(t_{i_2})\ldots V_\mathrm{I}(t_{i_n})\right] = \sum_{\lambda=1}^{n!} P_\lambda[V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)\ldots V_\mathrm{I}(t_n)\prod_{k=1}^{n-1}\Theta(t_k-t_{k+1})] \ , \quad (24.34)$$

avec $P_\lambda \in \mathcal{S}_n$. Cela étant l'opérateur $U_1(t)$ admet le développement :

$$\boxed{U_1(t) = \mathbf{1} + \sum_{n=1}^{+\infty} g^n \frac{(\mathrm{i}\hbar)^{-n}}{n!} \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \ldots \int_0^t \mathrm{d}t_n \, \mathcal{T}\left[V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)\ldots V_\mathrm{I}(t_n)\right]} \quad (24.35)$$

ce que l'on écrit parfois comme suit, pour faire court :

$$U_{\mathrm{I}}(t) = \mathcal{T}\,\mathrm{e}^{\frac{g}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t'\,V_{\mathrm{I}}(t')} \ , \tag{24.36}$$

en jouant symboliquement avec ce qui *ressemble* au développement en série entière de l'exponentielle (l'intégrale n-uple *n'est pas* une puissance n^{e}). Feynman [191] a proposé une autre technique pour respecter la chronologie dans l'exponentielle de $U(t)$.

Une écriture ou une autre du développement de $U(t)$ étant choisie pour la commodité du moment, l'amplitude de probabilité de la transition $\mathrm{i} \rightarrow \mathrm{f}$, $A_{\mathrm{i}\rightarrow\mathrm{f}}(t)$, (24.3), s'écrit alors :

$$A_{\mathrm{i}\rightarrow\mathrm{f}}(t) = \langle\psi_{\mathrm{f}}|U_0(t)\Big[\mathbf{1} + \sum_{n=1}^{+\infty} U_{\mathrm{I}}^{(n)}\Big]|\psi_{\mathrm{i}}\rangle \equiv \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}E_{\mathrm{f}}t}\, a_{\mathrm{i}\rightarrow\mathrm{f}}(t) \ , \tag{24.37}$$

l'écriture de droite venant de l'action de $U_0(t)$ sur le *bra* ; par définition :

$$a_{\mathrm{i}\rightarrow\mathrm{f}}(t) \stackrel{\text{déf}}{=} \delta_{\mathrm{if}} + \sum_{n=1}^{+\infty}\langle\psi_{\mathrm{f}}|U_{\mathrm{I}}^{(n)}|\psi_{\mathrm{i}}\rangle \equiv \sum_{n=0}^{+\infty} a_{\mathrm{i}\rightarrow\mathrm{f}}^{(n)}(t) \ ; \tag{24.38}$$

on note que $|A_{\mathrm{i}\rightarrow\mathrm{f}}(t)| = |a_{\mathrm{i}\rightarrow\mathrm{f}}(t)|$. Selon (24.27), on a alors successivement :

- à l'ordre zéro :

$$\boxed{a_{\mathrm{i}\rightarrow\mathrm{f}}^{(0)} = \delta_{\mathrm{if}}} \tag{24.39}$$

  Comme il se doit, l'amplitude de transition à l'ordre zéro, à partir d'un état propre de H_0, est nulle en l'absence de perturbation.

- au premier ordre :

$$a_{\mathrm{i}\rightarrow\mathrm{f}}^{(1)} = \frac{g}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t_1\,\langle\psi_{\mathrm{f}}|V_{\mathrm{I}}(t_1)|\psi_{\mathrm{i}}\rangle \ . \tag{24.40}$$

  Dans le cas où $|\psi_{\mathrm{i}}\rangle$ et $|\psi_{\mathrm{f}}\rangle$ sont des états propres[15] de H_0, l'action de U_0 contenu dans V_{I} peut être explicitée. En introduisant les pulsations :

$$\omega_{\mathrm{i}} = \hbar^{-1}E_{\mathrm{i}} \ , \ \omega_{\mathrm{f}} = \hbar^{-1}E_{\mathrm{f}} \ , \ \omega_k = \hbar^{-1}E_k \ , \ \omega_{\mathrm{fi}} = \omega_{\mathrm{f}} - \omega_{\mathrm{i}} \ , \ \omega_{kl} = \omega_k - \omega_l \ , \tag{24.41}$$

  on trouve :

$$\boxed{a_{\mathrm{i}\rightarrow\mathrm{f}}^{(1)} = \frac{g}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t_1\,\mathrm{e}^{\mathrm{i}\omega_{\mathrm{fi}}t_1}\langle\psi_{\mathrm{f}}|V(t_1)|\psi_{\mathrm{i}}\rangle} \tag{24.42}$$

- au second ordre :

$$a_{\mathrm{i}\rightarrow\mathrm{f}}^{(2)} = \Big(\frac{g}{\mathrm{i}\hbar}\Big)^2 \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\,\langle\psi_{\mathrm{f}}|V_{\mathrm{I}}(t_1)V_{\mathrm{I}}(t_2)|\psi_{\mathrm{i}}\rangle \ , \tag{24.43}$$

[15] C'est, de loin, le cas le plus fréquent.

soit :

$$a^{(2)}_{\mathrm{i}\to\mathrm{f}} = \left(\frac{g}{\mathrm{i}\hbar}\right)^2 \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \, \langle\psi_\mathrm{f}|U_0^\dagger(t_1)V(t_1)U_0(t_1)U_0^\dagger(t_2)V(t_2)U_0(t_2)|\psi_\mathrm{i}\rangle \quad ; \tag{24.44}$$

utilisant la propriété de groupe de U_0, il vient :

$$a^{(2)}_{\mathrm{i}\to\mathrm{f}} = \left(\frac{g}{\mathrm{i}\hbar}\right)^2 \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \, \langle\psi_\mathrm{f}|U_0^\dagger(t_1)V(t_1)U_0(t_1-t_2)V(t_2)U_0(t_2)|\psi_\mathrm{i}\rangle \quad ; \tag{24.45}$$

introduisant maintenant la décomposition de l'unité $\sum_k |E_k\rangle\langle E_k|$ juste à la droite de $V(t_1)$ (ou juste à gauche de $V(t_2)$), ou en écrivant explicitement :

$$U_0(t_1-t_2) = \sum_k |E_k\rangle \,\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\,E_k(t_1-t_2)}\langle E_k| \ , \tag{24.46}$$

on obtient :

$$\boxed{a^{(2)}_{\mathrm{i}\to\mathrm{f}} = \left(\frac{g}{\mathrm{i}\hbar}\right)^2 \sum_k \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \, \mathrm{e}^{\mathrm{i}\omega_{\mathrm{f}k}t_1}\,\mathrm{e}^{\mathrm{i}\omega_{k\mathrm{i}}t_2}\langle\psi_\mathrm{f}|V(t_1)|E_k\rangle\langle E_k|V(t_2)|\psi_\mathrm{i}\rangle} \tag{24.47}$$

et ainsi de suite. Sur un plan mnémotechnique, on remarque que l'ordre des indices des pulsations est le reflet exact de celui des éléments de matrice de la perturbation.

En résumé, les contributions successives à l'amplitude (24.3) $A_{\mathrm{i}\to\mathrm{f}}$ sont égales à :

$$A^{(0)}_{\mathrm{i}\to\mathrm{f}} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\,E_\mathrm{f}t}\delta_{\mathrm{if}} \ , \tag{24.48}$$

$$A^{(1)}_{\mathrm{i}\to\mathrm{f}} = g\frac{\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\,E_\mathrm{f}t}}{\mathrm{i}\hbar} \int_0^t \mathrm{d}t_1\,\mathrm{e}^{\mathrm{i}\omega_{\mathrm{fi}}t_1}\langle\psi_\mathrm{f}|V(t_1)|\psi_\mathrm{i}\rangle \ , \tag{24.49}$$

$$A^{(2)}_{\mathrm{i}\to\mathrm{f}} = g^2\frac{\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\,E_\mathrm{f}t}}{(\mathrm{i}\hbar)^2} \sum_k \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\,\mathrm{e}^{\mathrm{i}\omega_{\mathrm{f}k}t_1}\,\mathrm{e}^{\mathrm{i}\omega_{k\mathrm{i}}t_2}\langle\psi_\mathrm{f}|V(t_1)|E_k\rangle\langle E_k|V(t_2)|\psi_\mathrm{i}\rangle \ . \tag{24.50}$$

Il faut se souvenir que tous ces résultats supposent que les vecteurs $|\psi_\mathrm{i}\rangle$ et $|\psi_\mathrm{f}\rangle$ sont des états propres du Hamiltonien non-perturbé H_0, et définit le cadre naturel de raisonnement dans les situations physiques usuelles.

Les contributions aux ordres successifs peuvent être représentées par un diagramme comme montré sur la fig. 24.1. Chaque interaction est signalée par le sommet d'une ligne en pointillés. Pour le premier ordre, il suffit de sommer l'instant intermédiaire t_1 entre les deux temps fixés 0 et t. Au contraire, dès le second ordre, outre les intégrations sur les temps intermédiaires t_1 et t_2, apparaissent des sommations sur *tous* les états du système. Tous ces états, dits *virtuels*, sont à l'œuvre entre les deux instants variables t_1 et t_2 (avec la contrainte $t_2 < t_1$). Leur *virtualité* tient au fait qu'ils ne sont pas observables dans le processus dont ils contribuent à l'amplitude – quoiqu'ils sont états

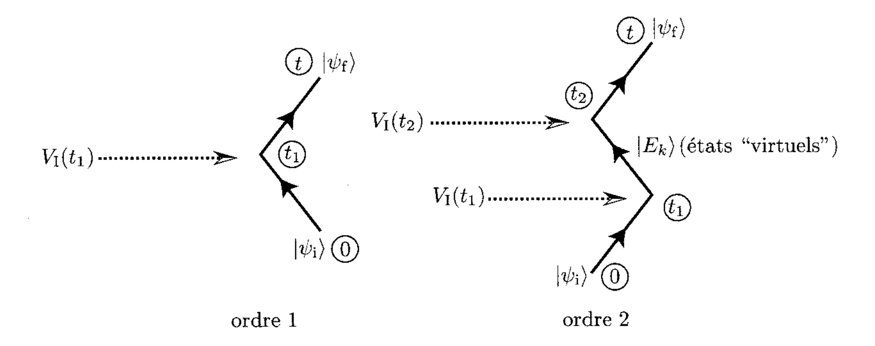


Figure 24.1: Illustration du mode de construction des amplitudes $A^{(1)}_{i\to f}$ et $A^{(2)}_{i\to f}$.

propres du système libre. Leur apparition dans le formalisme fait suspecter une violation de la conservation de l'énergie, mais bien sûr il n'en est rien. De telles représentations graphiques de la série de perturbation, et leurs généralisations, portent génériquement le nom de diagrammes de Feynman (pour une introduction, voir le livre de Mattuck [181]).

Les résultats précédents fournissent l'expression de la probabilité de transition $P_{i\to f}(t)$, $i \neq f$ entre deux états propres de H_0. S'en tenant au premier ordre, on a :

$$\boxed{P^{(1)}_{i\to f}(t) = g^2\hbar^{-2}\left|\int_0^t \mathrm{d}t_1\,\mathrm{e}^{\mathrm{i}\omega_{fi}t_1}\langle\psi_f|V(t_1)|\psi_i\rangle\right|^2} \tag{24.51}$$

À cet ordre, on a $P^{(1)}_{i\to f} = P^{(1)}_{f\to i}$, une symétrie qui ne tient plus aux ordres supérieurs. Toujours à cet ordre, la probabilité $P^{(1)}_{i\to f}$ est finalement donnée par le module carré de la transformée de Fourier de la perturbation vue par le système depuis le début jusqu'à l'instant t, transformée qui est prise à la valeur de la pulsation de Bohr associée aux deux états propres considérés.

Il n'est pas aisé de discuter en général la validité de l'approximation du premier ordre. Il est clair de toute façon que s'agissant en fait du deuxième terme d'un développement en série[16], et puisque la perturbation est réputée *petite*, les probabilités de transition doivent également être petites, soit $P_{i\to f} \ll 1$ à tout ordre, et notamment $P^{(1)}_{i\to f} \ll 1$. En particulier, il faut que l'intervalle de temps T pendant lequel la perturbation est appliquée ne soit pas trop long ; intuitivement, on attend une condition du genre :

$$g|\langle\psi_f|\bar{V}|\psi_i\rangle|\,T \ll \hbar \ , \tag{24.52}$$

où $\bar{V}$ est l'ordre de grandeur de la perturbation. Ceci se résume schématiquement en disant : la perturbation doit être assez faible et/ou de courte durée.

[16]Le premier terme est le terme d'ordre zéro, égal à δ_{if}.

♦ *Remarque*

D'une façon générale, la série de perturbation tronquée à un ordre donné ne fournit pas un opérateur d'évolution *unitaire* ; en pratique, ce défaut n'est pas rédhibitoire si les probabilités de transition sont très petites (c'est presque une nécessité méthodologique), mais il interdit d'établir formellement des propriétés générales exploitant des lois de conservation. Il existe plusieurs techniques préservant l'unitarité à tout ordre de la série tronquée. L'une d'entre elles est due à Pechukas et Light [192] ; elle consiste à poser :

$$U_1(t) = \mathrm{e}^{\mathrm{i}[gA_1(t)+g^2A_2(t)+\ldots]} \ , \tag{24.53}$$

et à écrire les équations récurrentes pour les opérateurs hermitiques $A_n(t)$ afin de déterminer ceux-ci. Il est clair que cette méthode *resomme* à l'infini une classe de processus élémentaires puisque, formellement, le développement de l'exponentielle tronquée $\mathrm{e}^{\mathrm{i}[gA_1(t)+g^2A_2(t)+\ldots+g^NA_N(t)]}$ donne des puissances arbitrairement grandes de g. Cette technique a été utilisée dans le problème 16.6.7 du Tome I.

D'autres formalismes sont disponibles (pour une présentation claire et synthétique, voir l'article de Wilcox [193]), utilisant notamment la structure d'algèbre de Lie des opérateurs impliqués, ainsi que le théorème de Cayley - Hamilton pour les systèmes possédant un nombre *fini* N d'états (un moment cinétique, un atome à N niveaux, etc.). L'esprit de ces approches est plutôt d'écrire l'opérateur d'évolution sous la forme d'un *produit* [194] :

$$U_1(t) = \mathrm{e}^{\mathrm{i}\Omega_1(t)}\mathrm{e}^{\mathrm{i}\Omega_2(t)}\ldots \ , \tag{24.54}$$

soit en décrétant que chaque opérateur hermitique $\Omega_k(t)$ est d'ordre g^k, soit en le développant en série entière de g.

Une autre grande méthode passe par la résolvante à laquelle on impose d'emblée des propriétés d'analyticité assurant l'unitarité (voir par exemple l'approximation dite *du pôle* dans la section 25.2, p. 1164). Ce procédé, lui aussi, resomme une classe de diagrammes de Feynman. ♦

24.3 Amplitudes de transition entre états discrets : exemples

Ce paragraphe est consacré à l'application des résultats précédents à deux cas concrets très importants ; on se bornera à ne retenir que le premier ordre, et seules seront considérées des transitions entre états bien séparés en énergie. Le cas où les états finals possibles sont au contraire très serrés fera l'objet d'un paragraphe particulier où sera obtenu le résultat important dit *Règle d'or de Fermi*.

24.3.1 Perturbation constante

Cette situation est celle où la perturbation est nulle jusqu'à un certain instant (pris comme origine) et est constante après. $V(t)$ contient en facteur la fonction échelon-unité $\theta(t)$ et peut s'écrire $V(t) = g\theta(t)V$, V étant un opérateur *indépendant* du temps. Alors, l'intégration en temps est immédiate ; partant de (24.49), il vient :

$$|A_{\mathrm{i}\rightarrow\mathrm{f}}^{(1)}| = g\hbar^{-1} \left|\langle\psi_{\mathrm{f}}|V|\psi_{\mathrm{i}}\rangle \int_0^t \mathrm{d}t_1\, \mathrm{e}^{\mathrm{i}\omega_{\mathrm{fi}}t_1}\right| \ . \tag{24.55}$$

Posant $\langle\psi_{\mathrm{f}}|V|\psi_{\mathrm{i}}\rangle = V_{\mathrm{fi}}$, on trouve :

$$|A_{\mathrm{i}\rightarrow\mathrm{f}}^{(1)}| = g\hbar^{-1}\, |V_{\mathrm{fi}}| \left|\frac{\mathrm{e}^{\mathrm{i}\omega_{\mathrm{fi}}t} - 1}{\mathrm{i}\omega_{\mathrm{fi}}}\right| \ , \tag{24.56}$$

d'où la probabilité, au premier ordre :

$$\boxed{P_{\mathrm{i}\rightarrow\mathrm{f}}^{(1)}(t) = g^2\hbar^{-2}\, |V_{\mathrm{fi}}|^2\, f(t,\, \omega_{\mathrm{fi}}) \qquad f(T,\, \omega) \stackrel{\text{déf}}{=} \frac{4}{\omega^2}\, \sin^2 \frac{\omega T}{2}} \tag{24.57}$$

La fonction $f(T,\, \omega)$ est homogène au carré d'un temps et assure que la probabilité est un nombre pur, comme il se doit. Elle a l'allure indiquée sur la figure 24.2, présentant un maximum – très aigu si $\omega T \gg 1$ – pour $\omega = 0$ où elle vaut T^2, et des maxima secondaires pour les valeurs de ω de la forme entier $\times\, \frac{2\pi}{T}$; sa largeur est d'ordre T^{-1} et son enveloppe décroît algébriquement $\propto \omega^{-2}$. En outre, il est facile de montrer par résidus que :

$$\int_{-\infty}^{+\infty} f(T,\, \omega)\, \mathrm{d}\omega = 2\pi T \ ; \tag{24.58}$$

en définitive, f joue le rôle d'un filtre.

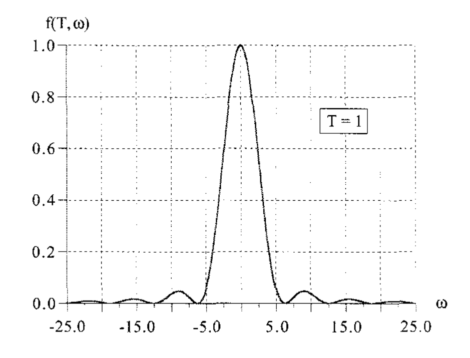


Figure 24.2: Graphe de la fonction $f(t, \omega)$ définie en (24.57).

Supposons que la perturbation dure pendant un temps T donné ; une fois qu'elle a cessé, les probabilités de transition entre états propres de H_0 deviennent constantes[17] :

$$P_{\mathrm{i}\rightarrow\mathrm{f}}(t \geq T) = P_{\mathrm{i}\rightarrow\mathrm{f}}(T) \simeq P_{\mathrm{i}\rightarrow\mathrm{f}}^{(1)}(T) = g^2\hbar^{-2}\, |V_{\mathrm{fi}}|^2\, f(T,\, \omega_{\mathrm{fi}}) \ . \tag{24.59}$$

[17] Alors le système évolue sous l'action de H_0 seul.

La propriété de filtrage de la fonction $f(T, \omega_{\rm fi})$ entraîne que, sous l'effet d'une perturbation constante[18], la probabilité de transition est maximale pour deux états de *même* énergie ($\omega_{\rm fi} = 0 \Longleftrightarrow E_{\rm f} = E_{\rm i}$) : pour un champ statique, les transitions se produisent majoritairement entre états dégénérés, ou quasi-dégénérés, à l'intérieur d'une bande d'énergie $\delta E \sim \hbar/T$. Pour les états d'énergie voisine, tels que $|\omega_{\rm fi}|T \ll 1$, la probabilité au premier ordre vaut à peu près :

$$P_{\rm i \rightarrow f}^{(1)}(T) \simeq g^2\hbar^{-2}\,|V_{\rm fi}|^2\,\frac{4}{\omega_{\rm fi}^2}\left(\frac{\omega_{\rm fi}T}{2}\right)^2 = g^2\hbar^{-2}\,|V_{\rm fi}|^2\,T^2\ , \tag{24.60}$$

et varie en temps comme le *carré* de la durée de la perturbation[19] ; cette variation est typique des perturbations de très courte durée entre états discrets[20]. L'expression (24.60) implique la valeur exacte de la fonction f à son maximum ; s'agissant d'une probabilité, ce résultat n'a de sens que dans la mesure où $g\hbar^{-1}|V_{\rm fi}|T \ll 1$. Il s'agit aussi d'une condition *nécessaire* pour que les termes suivants de la série de perturbation puissent être négligés ; cette condition confirme l'affirmation exprimée qualitativement par (24.52).

Enfin, on note que la largeur de la bande d'états concernés par la transition à partir de $|\psi_{\rm i}\rangle$ est évidemment reliée aux propriétés de dualité de la transformation de Fourier. La bande d'états impliqués autour de $\omega_{\rm i}$ a une largeur d'ordre $1/T$: elle est très large aux temps courts puisque, à ces instants, la perturbation est vue comme une fonction variant très vite et possède donc un spectre de Fourier très large. À l'inverse, pour $T \gg 1$, la perturbation est vue comme une fonction constante et a donc un spectre qui, sur l'échelle de fréquence du système, est très pointu, centré sur la fréquence nulle.

24.3.2 Perturbation sinusoïdale

Considérons maintenant le cas où $V(t)$ a l'une des deux formes :

$$V(t) = V\sin\omega t\ , \qquad V(t) = V\cos\omega t \qquad (\omega > 0)\ , \tag{24.63}$$

[18]à ceci près qu'elle est allumée à partir d'un certain instant pris comme origine.

[19]On ne peut donc pas en la circonstance définir une probabilité de transition par unité de temps qui est indépendante du temps.

[20]Pour tous les couples d'états bien séparés en énergie (tels que $\omega_{kl}t \gg 1$), f est quasiment nulle. En quelque sorte, la fonction f découple tout état initial donné $|\psi_{\rm i}\rangle$ de tous les autres états éloignés en énergie et semble se comporter à la manière d'une fonction de Dirac :

$$\forall\omega_k,\ \forall t \gg |\omega_{\rm i} - \omega_k|^{-1}\ :\ f(t, \omega_{k\rm i}) \simeq 2\pi t\,\delta(\omega_{k\rm i})\ ; \tag{24.61}$$

dans ces conditions, on est tenté d'en déduire :

$$P_{\rm i \rightarrow f}(t) \simeq 2\pi g^2\hbar^{-2}|V_{\rm fi}|^2\,t\,\delta(\omega_{\rm fi})\ . \tag{24.62}$$

En réalité, l'écriture à l'aide de la fonction δ ne s'impose nullement ici puisqu'il n'y a pas lieu à ce stade d'effectuer des intégrations en pulsation, et est plutôt source de confusion : elle donne l'impression (fausse) que $P_{\rm i \rightarrow f}^{(1)}(t) \propto t$. L'utilisation d'une fonction de Dirac présuppose l'intervention ultérieure dans une intégrale, mais ici, par hypothèse (états *discrets* bien séparés), aucune intégration n'est à prévoir. Si l'on y tient vraiment, on peut remplacer la fonction f par un *symbole de Kronecker*, mais cela n'apporte pas grand'chose par rapport à l'expression (24.57), qui est, elle, dénuée de toute ambiguïté.

où V est un opérateur indépendant du temps ; il est toujours licite de convenir que la pulsation ω a un signe donné, positif par exemple. La situation discutée ici modélise par exemple un atome à niveaux bien séparés couplé à un champ semi-classique monochromatique de pulsation $\omega > 0$. Partant de (24.55) et en posant $\langle\psi_{\mathrm{f}}|V|\psi_{\mathrm{i}}\rangle = V_{\mathrm{fi}}$, on a, pour le sinus :

$$|A^{(1)}_{\mathrm{i}\to\mathrm{f}}| = g\hbar^{-1}\,|V_{\mathrm{fi}}|\left|\frac{1}{2\mathrm{i}}\int_0^t \mathrm{d}t_1\,\mathrm{e}^{\,\mathrm{i}\omega_{\mathrm{fi}}t_1}(\mathrm{e}^{\,\mathrm{i}\omega t_1} - \mathrm{e}^{-\mathrm{i}\omega t_1})\right| \ , \tag{24.64}$$

d'où, au premier ordre :

$$P^{(1)}_{\mathrm{i}\to\mathrm{f}}(t) = g^2\frac{|V_{\mathrm{fi}}|^2}{4\hbar^2}\left|\frac{\mathrm{e}^{\mathrm{i}(\omega_{\mathrm{fi}}+\omega)t}-1}{\mathrm{i}(\omega_{\mathrm{fi}}+\omega)} - \frac{\mathrm{e}^{\mathrm{i}(\omega_{\mathrm{fi}}-\omega)t}-1}{\mathrm{i}(\omega_{\mathrm{fi}}-\omega)}\right|^2 . \tag{24.65}$$

Avec le cosinus, on obtient la même expression – au signe central près[21] :

$$P^{(1)}_{\mathrm{i}\to\mathrm{f}}(t) = g^2\frac{|V_{\mathrm{fi}}|^2}{4\hbar^2}\left|\frac{\mathrm{e}^{\mathrm{i}(\omega_{\mathrm{fi}}+\omega)t}-1}{\mathrm{i}(\omega_{\mathrm{fi}}+\omega)} + \frac{\mathrm{e}^{\mathrm{i}(\omega_{\mathrm{fi}}-\omega)t}-1}{\mathrm{i}(\omega_{\mathrm{fi}}-\omega)}\right|^2 . \tag{24.66}$$

Le signe $\pm$ dans (24.65) et (24.66) ne joue que pour les termes croisés venant du développement du module carré ; ce double produit est toujours très petit si on se place dans le cas où $\omega_{\mathrm{fi}}t \gg 1$; en effet, le module carré de chacun des deux termes contribuant à l'amplitude est une fonction du genre $f(t,\omega\pm\omega_{\mathrm{fi}})$, centrée en $\pm\omega_{\mathrm{fi}}$ et de largeur $\Delta\omega \sim 1/t$. Dans la limite $\omega_{\mathrm{fi}} \gg \Delta\omega$, le module au carré de l'amplitude se réduit quasiment à la somme des modules au carré de chaque contribution ; de la sorte, pour le sinus comme pour le cosinus on obtient la même expression approchée de la probabilité :

$$P^{(1)}_{\mathrm{i}\to\mathrm{f}}(t) \simeq g^2\frac{|V_{\mathrm{fi}}|^2}{4\hbar^2}\left[\left|\frac{\mathrm{e}^{\mathrm{i}(\omega_{\mathrm{fi}}+\omega)t}-1}{\mathrm{i}(\omega_{\mathrm{fi}}+\omega)}\right|^2 + \left|\frac{\mathrm{e}^{\mathrm{i}(\omega_{\mathrm{fi}}-\omega)t}-1}{\mathrm{i}(\omega_{\mathrm{fi}}-\omega)}\right|^2\right] . \tag{24.67}$$

soit :

$$\boxed{P^{(1)}_{\mathrm{i}\to\mathrm{f}}(t) \simeq g^2\frac{|V_{\mathrm{fi}}|^2}{4\hbar^2}\,[f(t,\omega+\omega_{\mathrm{fi}}) + f(t,\omega-\omega_{\mathrm{fi}})]} \tag{24.68}$$

où $f(t,\,\omega)$ est toujours la fonction définie en (24.57). Ainsi, et à condition que $t \gg \omega_{\mathrm{fi}}^{-1}$, la probabilité de transition présente ainsi un maximum très pointu pour les deux valeurs de la pulsation ω, conventionnellement prise positive : $\omega = \omega_{\mathrm{fi}}$ ou $\omega = -\,\omega_{\mathrm{fi}}$. Autrement dit, la probabilité est *résonnante* pour les états d'énergie E_{f} telle que :

$$E_{\mathrm{f}} = E_{\mathrm{i}} + \hbar\omega \qquad \text{ou} \qquad E_{\mathrm{f}} = E_{\mathrm{i}} - \hbar\omega \ . \tag{24.69}$$

Le premier cas représente l'absorption, le second l'émission. Pour un couple d'états donnés, la résonance se produit lorsque la fréquence du champ est égale à la fréquence de Bohr associée à ce couple. Par ailleurs, pour un ω_{fi} donné et puisque ω est supposé positif, seul l'un des deux termes dans (24.68) est important : c'est le terme dit résonnant, l'autre étant appelé antirésonnant. Finalement, pour une pertubation harmonique de durée T :

$$P^{(1)}_{\mathrm{i}\to\mathrm{f}}(t \geq T) \simeq g^2\frac{|V_{\mathrm{fi}}|^2}{4\hbar^2} \times \begin{cases} f(T,\omega+\omega_{\mathrm{fi}}) & \text{si } \omega_{\mathrm{fi}} < 0 \\ f(T,\omega-\omega_{\mathrm{fi}}) & \text{si } \omega_{\mathrm{fi}} > 0 \end{cases} . \tag{24.70}$$

[21] En faisant $\omega = 0$ dans (24.66), on retrouve bien la probabilité pour une perturbation constante donnée par (24.57).

Cette approximation est correcte si la largeur $\Delta\omega \sim 1/T$ de chaque fonction f est très petite devant la distance entre les deux résonances, soit $2|\omega_{\text{fi}}|$. Comme seules les valeurs de ω proches de $|\omega_{\text{fi}}|$ sont pertinentes, ceci signifie au total que l'on doit avoir :

$$\boxed{T \gg \frac{1}{\omega} \sim \frac{1}{|\omega_{\text{fi}}|}} \tag{24.71}$$

Le sens physique de cette condition est clair : la résonance en absorption ou en émission ne se produit que si la durée de la perturbation est grande par rapport à la période du champ appliqué et si $\omega \sim \omega_{\text{fi}}$; en d'autres termes, le champ externe doit apparaître comme quasi-monochromatique au système – lequel doit avoir le temps de compter un grand nombre de périodes – *et* avoir la bonne fréquence. C'est seulement lorsque la largeur de Fourier de la perturbation est très petite devant la fréquence propre du système que celui-ci peut présenter un comportement très sélectif se traduisant par une résonance. Pour une perturbation très brève à la même échelle, la réponse du système est évidemment à peu près plate en fréquence.

Les calculs explicites ci-dessus ne concernent que le premier ordre, et impliquent de façon décisive l'élément de matrice V_{fi}. Si ce dernier est nul, la probabilité est nulle au premier ordre – d'où la notion de *règles de sélection* (voir section 24.5.1). En pareil cas, il est nécessaire d'aller à l'ordre suivant et d'utiliser l'expression (24.50), qui contient *deux* interactions et où, de ce fait, figure une somme interne sur *tous* les états du systèmes (processus *virtuels*). Il est bien clair que deux états $|\psi_{\text{i}}\rangle$ et $|\psi_{\text{f}}\rangle$ non-couplés au premier ordre parce que $V_{\text{fi}} = 0$, peuvent être couplés au second ordre : il suffit pour cela qu'il existe au moins *un* état $|E_{k_0}\rangle$ tel que les deux éléments de matrice $\langle E_{k_0}|V|\psi_{\text{i}}\rangle$ et $\langle\psi_{\text{f}}|V|E_{k_0}\rangle$ sont chacun différents de zéro ; dès lors le couplage entre $|\psi_{\text{i}}\rangle$ et $|\psi_{\text{f}}\rangle$ se fait *via* l'état $|E_{k_0}\rangle$ et la transition $|\psi_{\text{i}}\rangle \longrightarrow |\psi_{\text{f}}\rangle$ devient possible. Avec une onde quasi-monochromatique de pulsation ω, la probabilité sera sensiblement non-nulle seulement si[22] $2\omega \simeq \hbar^{-1}|E_{\text{i}} - E_{\text{f}}|$. On parle alors de transition à deux photons ; de façon imagée, on peut dire que ceux-ci se font la courte échelle pour faire passer le système de $|\psi_{\text{i}}\rangle$ vers $|\psi_{\text{f}}\rangle$ en transitant par l'état (virtuel) $|E_{k_0}\rangle$. Il s'agit vraiment d'une image : à aucune étape de ce scénario par pas élémentaires l'énergie n'est conservée (il faut seulement que l'énergie soit conservée entre l'état initial et l'état final, soit $|E_{\text{i}} - E_{\text{f}}| = 2\hbar\omega$) – en particulier, l'état intermédiaire peut avoir une énergie absolument quelconque (il n'est pas au milieu).

24.4 Amplitudes de transition entre un état discret et un continuum d'états finals. Règle d'or de Fermi

Les résultats précédents portent sur les transitions entre deux états discrets d'énergie parfaitement déterminée ou au moins bien séparés de leurs plus proches voisins[23]. Cette

[22] L'intervention formelle du *carré* de la perturbation fait apparaître la fréquence *double* 2ω (voir problème 24.7.4).

[23] Dans ces conditions, le champ quasi-monochromatique *isole* les deux niveaux en harmonie avec la fréquence excitatrice ; c'est une condition *nécessaire* qui permet le cas échéant – moyennant quelques précautions – de faire l'approximation dite de l'atome à deux niveaux.

situation n'est pas la seule possible, loin de là ; en pratique, il est aussi important de savoir traiter la situation où la perturbation peut conduire le système dans une partie continue – ou quasi-continue – du spectre d'énergie. De nombreuses situations physiques correspondent à ce cas ; par exemple, dans l'ionisation d'un atome, l'électron éjecté passe d'un état discret (lié) à un état du continuum (non-lié). L'émission spontanée d'un photon par un atome ou un noyau excité est un autre exemple de transition vers un continuum à partir d'un état discret, ou en tout cas d'énergie très bien définie[24, 25]. Une autre situation est celle de la conversion d'énergie au sein d'une grosse molécule : après absorption lumineuse dans l'UV ou l'optique, l'énergie stockée dans la molécule se dilue dans les degrés de liberté denses constitués par les vibrations du squelette moléculaire (couplées d'ailleurs à la rotation de la molécule, voir chapitre 28, section 28.3) ; on parle alors de *transition non-radiative* pour décrire le phénomène où l'énergie n'est pas récupérée sous forme d'une émission de lumière – exemple de *conversion interne* ; le problème 14.6.11 (Tome I) propose une modélisation de l'étape de dilution, tout en explicitant le scénario de l'émergence de l'irréversibilité dynamique. Dans tous ces cas, un état initial évolue *irréversiblement* pourvu que la densité d'états accessibles soit suffisamment grande, comparée aux échelles de temps pertinentes physiquement : dimensionnellement parlant, le temps de récurrence éventuel est une fonction croissante de $\hbar/\delta E$, où δE est l'intervalle d'énergie typique entre deux états voisins de la variété dense.

D'après les postulats, lorsqu'une variable (observable) Ω prend des valeurs continues $\omega \in \mathbb{R}$, on suppose[26] qu'il existe une *densité* de probabilité, $P(\omega)$, égale à la *dérivée* d'une certaine fonction de répartition $F(\omega)$, laquelle donne la probabilité d'observer une valeur inférieure à une valeur donnée :

$$\text{Prob}[\Omega \leq \omega] \stackrel{\text{déf}}{=} F(\omega) \ , \quad \exists\, P(\omega) \stackrel{\text{déf}}{=} \lim_{\delta\omega \to 0} \frac{1}{\delta\omega}[F(\omega + \delta\omega) - F(\omega)] \ . \tag{24.72}$$

Avec cette densité $P(\omega)$, il est possible de calculer des probabilités pour que la variable aléatoire prenne des valeurs situées dans un intervalle donné, aussi petit soit-il. Par exemple, si $\Psi(x,\, t)$ est la fonction d'onde ordinaire, $|\Psi(x,\, t)|^2$ est la densité de probabilité de présence ; alors, la probabilité $P_{[ab]}(t)$ de trouver à l'instant t la particule entre les deux points d'abscisses a et b est :

$$P_{[ab]}(t) = \int_a^b |\Psi(x,\, t)|^2 \,\mathrm{d}x \ . \tag{24.73}$$

C'est bien l'intégrale (c'est-à-dire une *somme*) du module carré – non le module carré de l'intégrale – représentant la probabilité de trouver la particule entre a et b, obtenue en faisant la somme d'événements exclusifs (la particule est dans $[a, a + \delta x]$ *ou* dans

[24]On sait qu'il existe une relation du genre $\tau\delta\omega \sim 1$ entre la durée de vie τ de l'état excité et l'incertitude en énergie $\hbar\delta\omega$. Dans un cadre plus général (pour les particules élémentaires par exemple), un pic bien prononcé dans un spectre d'énergie est appelé *résonance*.

[25]Dans la situation la plus simple, le *continuum* est formé des états tensoriels de l'atome au fondamental par ceux d'un photon d'impulsion $\vec{k}$ et d'énergie $\varepsilon_k = \hbar kc$. Si le champ est enfermé dans une boîte de taille arbitrairement grande, la quantification des modes du champ a lieu sur une échelle d'énergie réellement *infiniment* petite, puisque la boîte de confinement du rayonnement est aussi grande que l'on veut.

[26]La Mécanique quantique écarte la possibilité de fonctions de répartition singulières continues au sens de Lebesgue.

$[a+\delta x, a+2\delta x]$, *ou* dans $[a+2\delta x, a+3\delta x]$ ... *ou* dans $[b-\delta x, b]$). Le passage de la (somme des amplitudes)^2, à la somme des (amplitudes)^2 traduit l'impossibilité pratique d'une résolution infiniment grande de la variable (observable) mesurée. En définitive, c'est la bonne échelle physique qui conditionne la bonne écriture au plan technique.

De la même façon, pour mesurer l'impulsion $\vec{p}$ d'une particule, il faut se référer à la représentation-p ; la fonction d'onde correspondante, $\Phi(\vec{p}, t)$, permet de définir une densité de probabilité pour les mesures de l'impulsion, soit $|\Phi(\vec{p}, t)|^2$. Le détecteur utilisé pour la mesure a une ouverture forcément finie, et donne un signal à chaque fois que l'impulsion est dans un domaine $\mathcal{D}$ centré sur l'extrémité du vecteur $\vec{p}$ choisi. Par exemple, on peut définir $\mathcal{D}$ par un petit angle solide $\Delta\Omega$ autour d'une direction donnée, ce qui définit l'orientation du détecteur, et par un ΔE qui définit sa résolution en énergie autour d'une valeur moyenne nominale – ou, de façon équivalente, le module de l'impulsion. En tout état de cause, pour une valeur nominale p_{f}, la probabilité $P_{\vec{p}_{\mathrm{f}}}$ d'obtenir un signal sur le détecteur s'obtient en sommant (intégrant) dans le domaine $\mathcal{D}$:

$$P_{\mathcal{D}}(t) \equiv P_{\vec{p}_{\mathrm{f}}}(t) = \int_{\vec{p}-\vec{p}_{\mathrm{f}} \in \mathcal{D}} |\Phi(\vec{p}, t)|^2 \, \mathrm{d}^3 p \ . \tag{24.74}$$

$\mathcal{D}$ est un (petit) volume dans l'espace des impulsions, déterminé par le protocole expérimental, fixé dès que l'orientation du détecteur, son ouverture angulaire et sa fenêtre en énergie ont été choisies. Bien évidemment, l'intégrale sur le vecteur $\vec{p}$ peut être remplacée par une intégration sur sa direction et son module, ce dernier étant relié à la seule énergie cinétique $E = \vec{p}^2/2m$ pour une particule libre dans l'état final[27]. Autrement dit, on effectue un simple changement de variables en écrivant :

$$\mathrm{d}^3 p = p^2 \mathrm{d}p \mathrm{d}\Omega = (2mE)\, \mathrm{d}(2mE)^{1/2}\, \mathrm{d}\Omega = (2m^3 E)^{1/2}\, \mathrm{d}E \mathrm{d}\Omega \overset{\text{déf}}{=} \rho(E)\, \mathrm{d}E \mathrm{d}\Omega \ . \tag{24.75}$$

où la dernière égalité définit la quantité $\rho(E)$, qui s'appelle densité d'états en énergie ; $\rho(E)\mathrm{d}E$ est le nombre d'états dans l'intervalle $[E, E+\mathrm{d}E]$. Dès lors, la probabilité précédente prend la forme :

$$P_{\mathcal{D}}(t) \equiv P_{E_{\mathrm{f}},\, \Delta\Omega}(t) = \int_{|E-E_{\mathrm{f}}| \leq \Delta E,\, |\Omega-\Omega_{\mathrm{f}}| \leq \Delta\Omega} \rho(E) |\Phi(\vec{p}, t)|^2 \, \mathrm{d}E \, \mathrm{d}\Omega \ . \tag{24.76}$$

Ces considérations se généralisent au cas où les états finals sont propres d'un ECOC contenant, outre l'énergie, des observables à spectre continu collectivement notées α (ci-dessus, α se réduit aux angles précisant la direction de $\vec{p}_{\mathrm{p_f}}$). Dans ces conditions, en continuant à désigner le vecteur d'état par $|\Psi(t)\rangle$ – la fonction d'onde étant $\Psi(\alpha, E, t)$ soit $\langle \alpha, E|\Psi(t)\rangle$ –, la probabilité d'obtenir l'énergie E_{f} et les valeurs α_{f} à ΔE et $\Delta\alpha$ près peut toujours s'écrire comme l'intégrale d'un module carré :

$$P_{E_{\mathrm{f}}\, \alpha_{\mathrm{f}}}(t) = \int_{|E-E_{\mathrm{f}}| \leq \Delta E,\, |\alpha-\alpha_{\mathrm{f}}| \leq \Delta\alpha} \rho(\alpha, E) |\langle \alpha, E|\Psi(t)\rangle|^2 \, \mathrm{d}E \, \mathrm{d}\alpha \ ; \tag{24.77}$$

[27]Supposer que la particule est libre dans l'état final est une hypothèse physiquement acceptable. Le détecteur est toujours situé à une distance macroscopique, infiniment grande par rapport à l'échelle de longueur qui caractérise le système émissif. Même pour un potentiel à portée infinie (le potentiel coulombien nu par exemple), introduire les vrais états non-liés est en général inutile puisque, par rapport aux états libres, ils n'introduisent que des corrections sous-dominantes dans l'expression du courant, donc de la section efficace (voir chapitre 26, Remarque p. 1213).

$|\langle\alpha, E|\Psi(t)\rangle|^2$ joue le rôle d'une *densité* de probabilité de transition vers l'état $|\alpha, E\rangle$, cependant que $\rho(\alpha, E)$ assure le décompte les états de même α et E à $\mathrm{d}\alpha$ et $\mathrm{d}E$ près. Cela étant, le calcul de l'amplitude de transition, au premier ordre, entre un état initial discret $|\psi_\mathrm{i}\rangle$ et un état du continuum $|\alpha, E\rangle\rangle$ démarre comme antérieurement : alors, aucune hypothèse quant à la nature (discrète ou continue) des états impliqués n'a été faite – sauf qu'ils sont propres de H_0. Ainsi, dans les notations en cours, et pour une perturbation de durée T, il faut trouver $\langle\alpha, E|\Psi(T)\rangle$, soit $\langle\alpha, E|U(T)|\psi_\mathrm{i}\rangle$; au premier ordre, (24.49) donne :

$$|\langle\alpha,\, E|\Psi(T)\rangle| \simeq |A^{(1)}_{\mathrm{i}\to\mathrm{f}\equiv(\alpha,E)}(T)| = \hbar^{-1}\left|\int_0^T \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(E-E_\mathrm{i})t_1}\,\langle\alpha, E|V(t_1)|\psi_\mathrm{i}\rangle\,\mathrm{d}t_1\right| \;. \qquad (24.78)$$

Par exemple, pour une perturbation constante qui agit à partir de l'instant $t = 0$, on a :

$$|\langle\alpha,\, E|\Psi(T)\rangle| \simeq \hbar^{-1}\left|\langle\alpha, E|V|\psi_\mathrm{i}\rangle\int_0^T \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(E-E_\mathrm{i})t_1}\,\mathrm{d}t_1\right| \;. \qquad (24.79)$$

Il en résulte que la densité de probabilité de transition est :

$$\hbar^{-2}\,|\langle\alpha, E|V|\psi_\mathrm{i}\rangle|^2\left|\int_0^T \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(E-E_\mathrm{i})t_1}\,\mathrm{d}t_1\right|^2 \equiv |\langle\alpha, E|V|\psi_\mathrm{i}\rangle|^2\, f(T,\, \frac{E-E_\mathrm{i}}{\hbar}) \;, \qquad (24.80)$$

où la fonction f a été définie en (24.57). La probabilité de transition entre $|\psi_\mathrm{i}\rangle$ et le groupe d'états finals définis par les valeurs E_f et α_f à ΔE et $\Delta\alpha$ près est alors donnée par l'intégrale sommant sur toutes les valeurs de la fenêtre du détecteur :

$$P_{E_\mathrm{f}\alpha_\mathrm{f}}(T) = \hbar^{-2}\int_{|E-E_\mathrm{f}|\le\Delta E,\,|\alpha-\alpha_\mathrm{f}|\le\Delta\alpha} \rho(\alpha, E)|\langle\alpha, E|V|\psi_\mathrm{i}\rangle|^2 f(T, \frac{E-E_\mathrm{i}}{\hbar})\,\mathrm{d}E\mathrm{d}\alpha \;. \qquad (24.81)$$

La fonction f varie très rapidement au voisinage de E_i et son intégrale en énergie vaut $2\pi\hbar T$. À l'inverse, puisque l'on est en présence d'un continuum d'états finals, la densité ρ et l'élément de matrice de V varient lentement, comparés à f : c'est précisément ce qui différencie un *continuum* d'énergie d'un spectre discret, où au contraire la densité a l'allure d'un peigne (irrégulier) de Dirac[28], variant à toute vitesse ; en toute rigueur d'ailleurs, d'un point de vue physique, il n'existe pas de vrai *continuum* : il s'agit toujours de l'idéalisation d'une situation physique où les états sont si proches en énergie les uns des autres qu'il est exclu de les séparer en pratique[29]. On peut donc, sur un strict plan opératoire, remplacer f par une fonction de Dirac, soit précisément par[30] $2\pi\hbar T\delta(E-E_\mathrm{i})$.

---

[28]puisqu'il n'y a d'états que pour des valeurs infiniment précises de l'énergie.

[29]Voir par exemple le traitement ci-après de l'effet photoélectrique. La quantification de l'impulsion par application des conditions cycliques de Born - von Kármán introduit une *discrétisation* des énergies que l'on peut rendre aussi petite que l'on veut, comparée à la résolution du détecteur le plus performant conceptuellement possible.

[30]Dirac a précisément introduit sa "fonction" comme un objet purement opératoire : associé à une *fonction lentement variable à l'échelle de ce qui va être modélisé par un δ, il sélectionne* la valeur de la fonction à variation lente au point de concentration de δ.

Se souvenir que la fonction $f(T,\,\omega)$ est homogène au carré d'un temps, d'où la substitution :

$$f(T,\, E/\hbar) \to CT\delta(E/\hbar) = CT\hbar\delta(E) \;, \qquad (24.82)$$

où C est un nombre, égal à 2π (voir (24.58)).

En effet, dans un premier temps, on peut extraire ρ et l'élément de matrice de l'intégrale d'énergie, en les prenant à la valeur E_i où est centrée la fonction f :

$$P_{E_\mathrm{f}\,\alpha_\mathrm{f}}(T) \simeq \hbar^{-2}\int_{|\alpha-\alpha_\mathrm{f}|\leq\Delta\alpha} \rho(\alpha, E_\mathrm{i})\,|\langle\alpha, E_\mathrm{i}|V|\psi_\mathrm{i}\rangle|^2\,\mathrm{d}\alpha \int_{|E-E_\mathrm{f}|\leq\Delta E} f\big(T, \frac{E-E_\mathrm{i}}{\hbar}\big)\,\mathrm{d}E \ . \quad (24.83)$$

Si l'énergie E_f n'est pas égale à E_i, l'intégrale de droite est quasiment nulle (schématiquement : on intègre un Dirac en dehors de son point de concentration). Au contraire, si $|E_\mathrm{f}-E_\mathrm{i}| \lesssim \mathcal{O}(\hbar/T)$ et si la largeur du détecteur ΔE est grande par rapport à $\hbar/T$, on peut rejeter les bornes sur l'énergie en $\pm\infty$ et on a alors à peu près :

$$\int_{|E-E_\mathrm{f}|\leq\Delta E} f(T, \frac{E-E_\mathrm{i}}{\hbar})\,\mathrm{d}E \simeq \int_{-\infty}^{+\infty} \mathrm{d}E\, f\big(T, \frac{E-E_\mathrm{i}}{\hbar}\big) = 2\pi\hbar T \ , \quad (24.84)$$

d'où, dans ces conditions et avec $E_\mathrm{f} \simeq E_\mathrm{i}$:

$$P_{E_\mathrm{f}\,\alpha_\mathrm{f}}(t) \simeq \frac{2\pi}{\hbar}T\int_{|\alpha-\alpha_\mathrm{f}|\leq\Delta\alpha} \rho(\alpha, E_\mathrm{i})\,|\langle\alpha, E_\mathrm{i}|V|\psi_\mathrm{i}\rangle|^2\,\mathrm{d}\alpha \ . \quad (24.85)$$

Si $E_\mathrm{f}-E_\mathrm{i}$ est plus grand que $\hbar/T$, l'intégrale en énergie est pratiquement nulle ; dans la limite $T \to +\infty$, $P_{E_\mathrm{f}\,\alpha_\mathrm{f}}$ n'est non nulle que si $E_\mathrm{f} = E_\mathrm{i}$: ce n'est rien d'autre que l'expression de la conservation de l'énergie pour un processus réel.

De fait, l'expression (24.85) permet de définir une probabilité de transition par unité de temps, p, qui est cette fois indépendante du temps :

$$p_{E_\mathrm{f}\,\alpha_\mathrm{f}} \stackrel{\text{déf}}{=} \frac{P_{E_\mathrm{f}\,\alpha_\mathrm{f}}}{T} = \frac{2\pi}{\hbar}\int_{|\alpha-\alpha_\mathrm{f}|\leq\Delta\alpha} \rho(\alpha, E_\mathrm{i})\,|\langle\alpha, E_\mathrm{i}|V|\psi_\mathrm{i}\rangle|^2\,\mathrm{d}\alpha \ . \quad (24.86)$$

Si on ne se préoccupe pas de la valeur de α dans l'état final (si par exemple on mesure seulement l'énergie), il faut intégrer sur α sans restriction (i.e. sur toutes les valeurs possibles de α), d'où :

$$\boxed{p_{E_\mathrm{f}} = \frac{2\pi}{\hbar}\int_{\forall\alpha} \rho(\alpha, E_\mathrm{i})\,|\langle\alpha, E_\mathrm{i}|V|\psi_\mathrm{i}\rangle|^2\,\mathrm{d}\alpha} \quad (24.87)$$

En général, il n'y a aucune raison pour que, après intégration, le résultat soit de la forme : une densité effective ρ multipliée par un élément de matrice. D'un autre côté, dans le cas où l'état final est caractérisé par sa seule énergie, il n'y a pas de grandeur du type α, donc pas d'intégration en α. Il reste alors :

$$\boxed{p_{E_\mathrm{f}} = \frac{2\pi}{\hbar}\,\rho(E_\mathrm{i})\,|\langle E_\mathrm{i}|V|\psi_\mathrm{i}\rangle|^2 \qquad (E_\mathrm{f} = E_\mathrm{i})} \quad (24.88)$$

d'une part ; d'autre part, si la condition sur l'énergie n'est pas satisfaite, la probabilité de transition est nulle dans la limite $t\Delta E \gg \hbar$. Le résultat (24.88) porte le nom de *Règle d'or de Fermi* et décrit – sur une grande échelle de temps à préciser dans chaque cas – ce qui se passe quand un état discret est couplé à un continuum d'états finals par

une perturbation constante, "allumée" à partir d'un certain instant. Elle exprime la conservation de l'énergie : la probabilité est donnée par l'expression (24.88) quand l'état final a une énergie E_f qui est *égale* à celle de l'état de départ, E_i, et est nulle autrement ; en définitive, ceci peut se traduire en introduisant un *symbole* de Kronecker :

$$p_{E_\mathrm{f}} = \frac{2\pi}{\hbar}\,\rho(E_\mathrm{i})\,|\langle\alpha_\mathrm{f}, E_\mathrm{f}|V|\psi_\mathrm{i}\rangle|^2\,\delta_{E_\mathrm{f}\,E_\mathrm{i}} \ . \tag{24.89}$$

Bien évidemment, le caractère *isoénergétique* de la transition – exprimé par le symbole de Kronecker dans (24.89) – résulte du fait que la perturbation V a été prise *constante*. Avec une perturbation monochromatique de pulsation ω, on trouve $\delta_{E_\mathrm{f}\,E_\mathrm{i}\pm\hbar\omega}$ (voir ci-dessous le traitement de l'effet photoélectrique, sous-section 24.5.2).

Figure 24.3: Illustration de la Règle d'or de Fermi avec une perturbation constante.

24.5 Applications exemplaires

Il s'agit ici de donner deux applications du formalisme précédent dans deux situations physiques importantes et constituant deux exemples de transitions soit entre états discrets, soit entre un état discret et un *continuum*.

24.5.1 Interaction d'un atome avec un champ électromagnétique (description semi-classique)

On va calculer au premier ordre les probabilités de transition induites dans un atome (à un seul électron, de masse μ) par un champ électromagnétique considéré comme extérieur (donc non-quantifié) et représenté par une simple onde plane monochromatique polarisée ([25], $\mathrm{A}_{\mathrm{XIII}}$). Le but ultime est de montrer, le plus simplement possible, l'existence de règles de sélection gouvernant les intensités des raies des spectres atomiques.

Pour une telle onde plane, le potentiel scalaire peut toujours être pris égal à zéro (voir [25], Appendice III, 4-b-α) ; les deux champs sont :

$$\vec{\mathcal{E}}(\vec{r}, t) = \mathcal{E}\,\vec{a}\,\cos(\vec{k}.\vec{r} - \omega t) \ , \qquad \vec{\mathcal{B}}(\vec{r}, t) = \mathcal{B}\,\vec{b}\,\cos(\vec{k}.\vec{r} - \omega t) \ , \tag{24.90}$$

où $\vec{a}$ et $\vec{b}$ sont les polarisations, $(\vec{b}, \vec{n} = \vec{k}/k, \vec{a})$ est un repère orthonormé direct (en particulier $\vec{k}.\vec{a} = 0$) ; dans la suite, on choisit l'axe Ox le long de $\vec{b}$, Oz selon $\vec{a}$ et Oy le

long du vecteur d'onde $\vec{k}$. Les modules des champs sont reliés par l'égalité $\mathcal{E} = \mathcal{B}c$, c étant la vitesse de la lumière dans le vide. On vérifie immédiatement que le potentiel-vecteur :

$$\vec{A}(\vec{r},\, t) = A_0\, \vec{a}\, \sin(\vec{k}.\vec{r} - \omega t) \ , \qquad A_0 = \frac{\mathcal{E}}{\omega} = \frac{\mathcal{B}}{k} \tag{24.91}$$

restitue bien les deux champs (24.90) par les relations $\vec{\mathcal{E}} = -\partial_t \vec{A}$ et $\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$. Le Hamiltonien se forme comme d'habitude ; complété à la main par l'interaction entre le champ externe et le spin électronique – mais en négligeant les termes de structure fine, pour simplifier – il s'écrit :

$$H = \frac{1}{2\mu}\,(\vec{p} - e\vec{A})^2 + V(\vec{r}) - g_{\mathrm{e}}\,\frac{e}{2\mu}\,\vec{S}.\vec{\mathcal{B}} \ . \tag{24.92}$$

$V(\vec{r})$ est le potentiel intra-atomique et $g_{\mathrm{e}} \simeq 2$. En développant l'expression (24.92) et en négligeant le terme diamagnétique quadratique en $\vec{A}$, on trouve :

$$H = \left[\frac{\vec{p}^{\,2}}{2\mu} + V(\vec{r})\right] - \left[\frac{e}{2\mu}(\vec{A}.\vec{p} + \vec{p}.\vec{A}) + g_{\mathrm{e}}\,\frac{e}{2\mu}\,\vec{S}.\vec{\mathcal{B}}\right] \equiv H_0 + V(t) \ . \tag{24.93}$$

H_0 a pour vecteurs et énergies propres $|\psi_n\rangle$ et E_n. $V(t)$ est la somme de deux termes ; comme $\vec{\nabla}.\vec{A} \propto \vec{k}.\vec{a} = 0$, $\vec{A}.\vec{p} = \vec{p}.\vec{A}$ et on peut écrire :

$$V = V_1 + V_2 \ , \qquad V_1 \stackrel{\text{déf}}{=} -\frac{e}{\mu}\vec{A}.\vec{p} \ , \quad V_2 \stackrel{\text{déf}}{=} -g_{\mathrm{e}}\,\frac{e}{2\mu}\,\vec{S}.\vec{\mathcal{B}} \ . \tag{24.94}$$

Ces deux termes ne sont pas du même ordre de grandeur ; en effet($p \sim \hbar/a_0$) :

$$\frac{\langle V_2\rangle}{\langle V_1\rangle} \sim \frac{\langle\vec{S}\,\rangle.\vec{\mathcal{B}}}{\vec{A}.\langle\vec{p}\rangle} \sim \frac{\hbar\, kA_0}{A_0\, p} = \frac{\hbar k}{p} \sim ka_0 \tag{24.95}$$

Dans le domaine optique ($\lambda \sim 5\,000$ Å), $ka_0 \sim (1/1\,000) \times 0,5 \ll 1$; une première approximation consiste donc à délaisser le terme V_2, quitte à y revenir par la suite. V_1 peut d'ailleurs lui-même être simplifié ; en effet, d'après (24.91), on a :

$$V_1 = -\frac{e}{\mu}\, A_0\, \vec{p}.\vec{a}\, \sin(\vec{k}.\vec{r} - \omega t) \ . \tag{24.96}$$

Le calcul des éléments de matrice de V_1 implique des fonctions d'onde atomiques, l'intégrand ne prenant de valeurs significatives que pour $\|\vec{r}\| \lesssim a_0$; dans ces conditions, le produit ky venant du produit scalaire $\vec{k}.\vec{r}$ est toujours très petit devant 1 sur le domaine pertinent d'intégration (là où la fonction est sensiblement non-nulle), d'où :

$$V_1 \simeq +\frac{e}{\mu}\, A_0\, \vec{p}.\vec{a}\, \sin\omega t \stackrel{\text{déf}}{=} V_{\mathrm{E1}} \ . \tag{24.97}$$

Cette approximation dite *dipolaire électrique* (notée E1) revient notamment à négliger, dans le domaine optique, le déphasage spatial de l'onde sur l'espace atomique effectif.

♦ *Remarques*

1. Il est instructif d'écrire les équations de Heisenberg en présence de V_{E1} ; on trouve :

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{r}_{\mathrm{H}} = \frac{\vec{p}_{\mathrm{H}}}{\mu} + \frac{e\mathcal{E}}{\mu\omega}\vec{a}\,\sin\omega t\ , \qquad \frac{\mathrm{d}}{\mathrm{d}t}\vec{p}_{\mathrm{H}} = -\,(\vec{\nabla}V)_{\mathrm{H}}\ . \tag{24.98}$$

Ainsi, au mouvement intra-atomique se superpose une oscillation forcée le long du champ électrique appliqué.

2. La forme de l'interaction dipolaire électrique définie en (24.97) peut surprendre : on aurait plutôt attendu une expression du genre $V'_{\mathrm{E1}} = -\vec{d}.\vec{\mathcal{E}}$ où $\vec{d}$ est le moment dipolaire électrique. En fait, (24.97) résulte directement du choix de la jauge (24.91) (le potentiel scalaire étant par ailleurs pris égal à zéro). Un changement de jauge approprié donne, au même niveau d'approximation, l'expression V'_{E1}. ♦

Afin de calculer les amplitudes de transition dans l'approximation E1, il faut connaître les éléments de matrice de $\vec{a}.\vec{p} = p_z$ entre les états propres de H_0 :

$$\langle\psi_{\mathrm{f}}|V_{\mathrm{E1}}|\psi_{\mathrm{i}}\rangle = +\frac{e\mathcal{E}}{\mu\omega}\,\langle\psi_{\mathrm{f}}|\vec{a}.\vec{p}|\psi_{\mathrm{i}}\rangle\,\sin\omega t \;\equiv\; +\frac{e\mathcal{E}}{\mu\omega}\,\langle\psi_{\mathrm{f}}|p_z|\psi_{\mathrm{i}}\rangle\,\sin\omega t\ . \tag{24.99}$$

Comme précédemment, $|\psi_{\mathrm{i}}\rangle$ et $|\psi_{\mathrm{f}}\rangle$ désignent les états extrêmes de la transition et sont supposés être propres de H_0. En utilisant l'équation :

$$\mathrm{i}\hbar\,\frac{\vec{p}}{\mu} = [\vec{r},\, H_0]\ , \tag{24.100}$$

on obtient[31] :

$$\boxed{\langle\psi_{\mathrm{f}}|V_{\mathrm{E1}}|\psi_{\mathrm{i}}\rangle = \mathrm{i}e\mathcal{E}\,\frac{\omega_{\mathrm{fi}}}{\omega}\,\sin\omega t\;\langle\psi_{\mathrm{f}}|\vec{r}.\vec{a}|\psi_{\mathrm{i}}\rangle} \tag{24.101}$$

Finalement, les éléments de matrice de V_{E1} sont bien proportionnels à ceux du moment dipolaire électrique.

Pour que la transition $|\psi_{\mathrm{i}}\rangle \rightarrow |\psi_{\mathrm{f}}\rangle$ puisse se produire dans l'approximation E1, il est donc nécessaire et suffisant[32] que $\langle\psi_{\mathrm{f}}|\vec{r}.\vec{a}|\psi_{\mathrm{i}}\rangle$ soit différent de zéro. Pour un champ central, les états $|\psi_{\mathrm{i}}\rangle$ et $|\psi_{\mathrm{f}}\rangle$ sont de la forme $R_{nl}Y_{lm}$; les harmoniques sphériques étant de parité déterminée et $\vec{r}$ étant impair, une condition nécessaire est que $|\psi_{\mathrm{i}}\rangle$ et $|\psi_{\mathrm{f}}\rangle$ soient de parité *opposée*. Très précisément, on peut écrire les composantes de $\vec{r}$ à l'aide de Y_{1m} :

$$z = r\sqrt{\frac{4\pi}{3}}\,Y_{10}\ , \qquad x \pm \mathrm{i}y = \mp\, r\sqrt{\frac{8\pi}{3}}\,Y_{1\pm1}\ . \tag{24.102}$$

[31] Par ailleurs, l'atome absorbe ou émet quand $\omega \simeq |\omega_{\mathrm{fi}}|$ (résonance). En pratique, on récrit souvent l'expression (24.101) sous la forme simplifiée $\langle\psi_{\mathrm{f}}|V_{\mathrm{E1}}|\psi_{\mathrm{i}}\rangle \simeq \mathrm{i}e\mathcal{E}\,\sin\omega t\,\langle\psi_{\mathrm{f}}|\vec{r}.\vec{a}|\psi_{\mathrm{i}}\rangle$. $e\vec{r}$ étant le moment dipolaire électrique $\vec{d}$, $\langle\psi_{\mathrm{f}}|V_{\mathrm{E1}}|\psi_{\mathrm{i}}\rangle$ est finalement proportionnel à $\langle\psi_{\mathrm{f}}|\vec{d}|\psi_{\mathrm{i}}\rangle$, d'où la terminologie *transition dipolaire électrique*.

[32] tant que l'on reste au premier ordre de la théorie des perturbations.

Il en résulte que l'intégrale angulaire de $\langle\psi_\text{f}|V_\text{E1}|\psi_\text{i}\rangle$ contient un facteur de la forme :

$$\int Y^*_{l_\text{f}\,m_\text{f}}\, Y_{1m}\, Y_{l_\text{i}m_\text{i}}\,\mathrm{d}\Omega \qquad (m = 0,\, \pm 1)\ . \tag{24.103}$$

Le produit des deux Y de droite est un produit tensoriel d'états propres de deux moments cinétiques égaux l'un à 1, l'autre à l_i ; le moment somme peut donc prendre les valeurs entières $l_\text{i} - 1$ (si $l_\text{i} > 0$), l_i et $l_\text{f} + 1$. En tant que produit scalaire, l'intégrale est sûrement nulle si l_f n'est pas égale à l'une de ces trois valeurs ; par ailleurs, elle est nulle par parité si $l_\text{f} = l_\text{i}$. Au total, elle est nulle si $|l_\text{i} - l_\text{f}|$ n'est pas égal à 1, ou si $-m_\text{f} + 1 + m_\text{i}$ n'est pas nul. Ceci permet d'énoncer les règles de sélection dipolaires électriques donnant les symétries des états extrêmes pouvant donner lieu à ces transitions :

$$\boxed{\Delta l = \pm 1\ ,\quad \Delta m = 0,\, \pm 1 \qquad (\text{E1})} \tag{24.104}$$

Pour une polarisation linéaire du champ électrique, il faut $\Delta m = 0$; $\Delta m = \pm 1$ vaut pour les champs polarisés circulairement dans un plan perpendiculaire à l'axe de quantification. Les relations (24.104) étant satisfaites, les intensités relatives des raies permises dans l'approximation E1 sont (essentiellement) fixées relativement par la valeur de l'intégrale radiale $\int_{\mathbb{R}+} r^2 R_{n_\text{f}l_\text{f}}\, r\, R_{n_\text{i}l_\text{i}} \mathrm{d}r$.

Quand une transition est interdite dans l'approximation E1 parce que les règles (24.104) ne sont pas satisfaites, ceci ne signifie pas qu'elle est inobservable, mais seulement qu'elle est *a priori* de faible intensité. En pareil cas, il convient d'aller un cran plus loin dans les approximations ; par exemple – et si on décide de rester au premier ordre de la théorie des perturbations –, on peut introduire la première correction au déphasage spatial de l'onde en développant le sinus ; on trouve ainsi[33] :

$$V_1 \simeq V_\text{E1} - \frac{e\mathcal{B}}{\mu}\,(\vec{p}.\vec{a})\,(\vec{n}.\vec{r})\,\cos\omega t\ . \tag{24.105}$$

L'apparition du module du champ magnétique montre que la cohérence du calcul exige d'introduire simultanément le couplage direct entre le spin et ce champ, donnant lieu au terme V_2 ; en effet, d'après (24.105), l'écart à V_E1 est d'ordre $(e\mathcal{B}/\mu)pr \sim (e\mathcal{B}/\mu)\,\hbar$, alors que $V_2 \sim (e/\mu)\hbar\mathcal{B}$. Toujours pour la cohérence, il faut développer V_2 au premier ordre en $\vec{k}.\vec{r}$. Quelques manipulations simples sur le terme correctif dans (24.105) font notamment apparaître le moment cinétique orbital $\vec{L}$. Finalement, après regroupement de tous les termes, l'opérateur $V(t)$ complet (24.94) apparaît sous la forme approchée :

$$V(t) \simeq V_\text{E1} - \frac{e}{2\mu}(\vec{L} + g_\text{e}\vec{S}).\vec{\mathcal{B}}\cos\omega t - \frac{e}{2\mu\omega}[(\vec{k}.\vec{r})(\vec{\mathcal{E}}.\vec{p}) + (\vec{k}.\vec{p})(\vec{\mathcal{E}}.\vec{r})]\,\cos\omega t \equiv V_\text{E1} + V_\text{M1} + V_\text{E2}\ , \tag{24.106}$$

où $\vec{\mathcal{E}}$ et $\vec{\mathcal{B}}$ représentent ici simplement $\vec{\mathcal{E}}(\vec{0},\, t)$ et $\vec{\mathcal{B}}(\vec{0},\, t)$. Le terme contenant les moments cinétiques est appelé dipolaire magnétique[34], puisqu'il représente le couplage direct entre le champ magnétique et le moment (dipolaire) magnétique total de l'atome. Le

[33] Au passage, on utilise la relation $kA_0 = \mathcal{B}$. Par ailleurs, $\vec{n}$ désigne un vecteur unitaire le long de $\vec{k}$.
[34] Pour dipolaire magnétique, l'acronyme consacré est M1.

deuxième terme, V_{E2}, implique des éléments de matrice de combinaisons quadratiques des coordonnées[35] et est appelé pour cette raison couplage quadrupolaire électrique, symboliquement E2.

Il est maintenant possible d'énoncer de nouvelles règles de sélection. En ce qui concerne les transitions M1, pour qu'elles soient possibles, il faut d'abord $\Delta l = 0$ puisque $(\vec{L} + g_{\mathrm{e}}\vec{S})$ ne saurait changer le nombre quantique l. Il faut en outre $\Delta m_l = 0, \pm 1$ ou $\Delta m_s = 0, \pm 1$.

Pour les transitions E2, il faut bien sûr $\Delta m_s = 0$. En ce qui concerne l'intégrale angulaire $\langle Y_{l_{\mathrm{f}}m_{\mathrm{f}}}|Y_{2m}|Y_{l_{\mathrm{i}}m_{\mathrm{i}}}\rangle$, des arguments analogues à ceux utilisés plus haut pour les transitions E1 montrent que les transitions E2 ne sont pas interdites si :

$$\boxed{\Delta l = 0, \pm 2 \ , \quad \Delta m = 0, \pm 1, \pm 2 \qquad (\mathrm{E2})} \tag{24.107}$$

Ces considérations se généralisent à un ordre multipolaire arbitraire. D'un autre côté, il faut se souvenir que l'analyse théorique des raies atomiques contient deux approximations emboîtées : d'une part le développement multipolaire amorcé ci-dessus, d'autre part les ordres successifs de la théorie des perturbations. Dans toute situation précise, il convient de comparer entre elles les différentes contributions provenant de ces deux types de développements afin d'effectuer un calcul cohérent, et en relation avec la précision expérimentale. Le problème 24.7.4 donne l'idée des transitions à deux photons.

24.5.2 Théorie élémentaire de l'effet photoélectrique

Il s'agit de calculer, au premier ordre, la probabilité de transition entre l'état fondamental $|\psi_{\mathrm{i}}\rangle = |1\mathrm{s}\rangle \equiv |\psi_{100}\rangle$ et un état du *continuum* de l'atome d'hydrogène, sous l'effet d'un champ électromagnétique traité à nouveau classiquement [80]. Ceci constitue sans doute l'exemple le plus simple d'une transition entre un état discret (lié) et un état du *continuum* (non-lié), et constitue l'archétype de la fragmentation d'un système rendu instable par une action extérieure dynamique. Dans l'état final, l'électron n'est pas rigoureusement "libre" (il reste soumis au potentiel Coulombien), même s'il se trouve arbitrairement loin du noyau ; cependant, si l'énergie finale est grande à l'échelle de l'eV, on peut admettre que l'état Coulombien non-lié est en fait très voisin d'une onde plane. Selon Einstein, la réaction peut s'écrire :

$$\text{photon} + \text{électron lié} \ \rightarrow \ \text{électron non-lié} \ . \tag{24.108}$$

Le photon est absorbé par l'atome (il n'apparaît plus dans l'état final) ; en considérant l'électron à l'infini dans l'état final, la conservation de l'énergie se traduit par :

$$\hbar q c + E_1 = \frac{1}{2}\mu \vec{v}_{\mathrm{f}}^{\,2} \equiv E_{\mathrm{f}} \qquad (q = \|\vec{q}\|) \ , \tag{24.109}$$

[35] On s'en convainc en utilisant (24.100).

$\hbar\vec{q}$ est l'impulsion du photon incident, $\hbar qc \equiv \hbar\omega$ son énergie, $E_1 \simeq -13,6$ eV est l'énergie de l'électron quand il est lié dans l'état fondamental. Pour un électron dans un métal, le seul changement à effectuer dans (24.109) consiste à remplacer E_1 par $-W_S$, puisque le travail de sortie est traditionnellement défini comme une quantité positive.

On choisit à nouveau le potentiel-vecteur sous la forme (24.91) et on retient l'interaction la plus simple, V_1, donnée par (24.96), que l'on récrit sous la forme :

$$V_1(t) = w\,\mathrm{e}^{\mathrm{i}\omega t} + w^\dagger\,\mathrm{e}^{-\mathrm{i}\omega t} \quad , \quad w = \frac{e}{2\mathrm{i}\mu} A_0\,\vec{a}.\vec{p}\,\mathrm{e}^{-\mathrm{i}\vec{q}.\vec{r}} \quad , \quad w^\dagger = \frac{\mathrm{i}e}{2\mu} A_0\,\vec{a}.\vec{p}\,\mathrm{e}^{+\mathrm{i}\vec{q}.\vec{r}} \quad , \tag{24.110}$$

avec $\omega > 0$. En toute rigueur, l'amplitude de transition (à un ordre donné de la théorie de perturbation) s'exprime à l'aide des éléments de matrice entre l'état $|1\mathrm{s}\rangle$ et un état du *continuum* du potentiel Coulombien. Ces états sont connus (voir chapitre 19, (19.195)) ; afin de s'en tenir à une description élémentaire, on suppose que l'énergie des photons incidents est très grande devant l'énergie de liaison de l'électron, $|E_1|$; l'énergie finale de celui-ci est alors très voisine de $\hbar\omega$, $E_\mathrm{f} \simeq \hbar\omega$ et, comme anticipé ci-dessus, il est licite de prendre pour état final de l'électron, $|\psi_\mathrm{f}\rangle$, un état strictement libre, simple combinaison linéaire d'ondes planes dont le module du vecteur d'onde $\vec{k}$ est fixé par[36] :

$$E_\mathrm{f} = \frac{\hbar^2\vec{k}^{\,2}}{2\mu} \ . \tag{24.111}$$

La légitimité de ce choix repose sur le fait que les états non-liés *fortement excités* de l'atome d'hydrogène, s'ils ne sont pas strictement des ondes planes, doivent y ressembler fortement – pour des raisons physiques évidentes.

Le choix des ondes planes étant fait, une petite difficulté technique se présente, comme à chaque fois qu'il s'agit de manipuler des états non-normalisables. Pour traiter cette difficulté, différents trucs peuvent être utilisés. L'un d'entre eux consiste à imaginer que l'électron libéré reste toutefois confiné dans une boîte de dimension arbitrairement grande comparée aux autres longueurs déjà présentes dans le problème physique[37]. Il est clair que les prévisions physiques ne sauraient dépendre de la taille de cette boîte[38], supposée en général cubique pour simplifier, de côté L.

On a alors une particule dans un puits (infini) de taille linéaire L, dont on sait que les états sont quantifiés ; leur énergie s'exprime à l'aide de trois entiers strictement positifs n_x, n_y et n_z :

$$E_{n_x\,n_y\,n_z} = \frac{\hbar^2\vec{k}^{\,2}}{2\mu} = \frac{\hbar^2}{2\mu}(n_x^2 + n_y^2 + n_z^2)\frac{\pi^2}{L^2} \qquad (n_u = 1,\,2,\,3,\,\ldots) \ . \tag{24.112}$$

[36] Si l'énergie du photon est supposée grande devant E_i, on suppose néanmoins que le photoélectron n'est pas relativiste.

[37] Pour fixer les idées, cette boîte peut être choisie comme étant la pièce où se déroule l'expérience !

[38] Il peut arriver que la véracité de cette affirmation repose sur des subtilités mises ici sous le tapis, notamment quand il existe des interactions à longue portée comme par exemple l'interaction de Coulomb nue, qui décroît comme $1/r$. Alors, il n'est pas forcément facile de trouver une échelle de longueur propre au problème traité, ou, plus précisément, cette échelle peut être formellement infinie. En pratique, ou bien on tient compte d'effets physiques ignorés jusqu'alors – et fabriquant automatiquement des effets d'écran –, ou bien on introduit une fonction de coupure formelle (pour le Coulombien, on prend souvent un potentiel de Yukawa, $\mathrm{e}^{-r/\xi}/r$), et on prend astucieusement la limite (ξ infini, L infini).

L étant aussi grand que l'on veut par rapport à une distance typique du problème (ici, a_0 par exemple), il est bien clair que sur une échelle d'énergie atomique, ces énergies – formellement quantifiées – sont extrêmement rapprochées les unes des autres et constituent un quasi-*continuum* d'énergie[39]. Par ailleurs, pour le puits cubique infini situé dans l'octant positif de $\mathbb{R}^3$ et dont l'un des sommets est l'origine du repère, on sait que les états propres sont :

$$\psi_{n_x n_y n_z} = \left(\frac{2}{L}\right)^{3/2} \sin n_x \frac{\pi x}{L} \sin n_y \frac{\pi y}{L} \sin n_z \frac{\pi z}{L} \; . \tag{24.113}$$

La pratique montre que ces fonctions sont d'un usage peu agréable et conduisent à des calculs inutilement compliqués – puisque de toute façon, on s'attend à ce que ce qui se passe aux bords de la boîte de confinement n'ait aucune pertinence physique. Pour cette raison, on utilise le plus souvent des ondes planes[40] $\psi_{\vec{k}}(\vec{r}) = C\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$, définies dans une boîte cubique de côté L et auxquelles on impose d'autres conditions aux limites, dites de Born - von Kármán. Ces conditions s'énoncent comme suit :

$$\mathrm{e}^{-\mathrm{i}k_x L/2} = \mathrm{e}^{+\mathrm{i}k_x L/2} \; , \qquad \mathrm{e}^{-\mathrm{i}k_y L/2} = \mathrm{e}^{+\mathrm{i}k_y L/2} \; , \qquad \mathrm{e}^{-\mathrm{i}k_z L/2} = \mathrm{e}^{+\mathrm{i}k_z L/2} \; . \tag{24.114}$$

Chacune d'entre elles identifie un point d'abscisse $-L/2$ avec le point d'abscisse $+L/2$ situé sur la face opposée du cube, comme si les deux faces parallèles étaient juxtaposées (pour cette raison, les conditions (24.114) sont dites *cycliques*). À une dimension, ceci revient, au lieu de considérer un grand segment de droite s'étendant entre les abscisses $\pm L/2$, à refermer ce segment en boucle ; si L est assez grand devant toutes les autres longueurs, la courbure artificielle ainsi introduite est sans effet.

Il résulte immédiatement de (24.114) qu'il existe trois entiers quelconques (pris maintenant dans $\mathbb{Z}$ et non plus dans $\mathbb{N}$) donnant les composantes (quantifiées !) du vecteur $\vec{k}$:

$$k_x = n_x \frac{2\pi}{L} \; , \qquad k_y = n_y \frac{2\pi}{L} \; , \qquad k_z = n_z \frac{2\pi}{L} \; , \tag{24.115}$$

de sorte que l'espace du vecteur d'onde est discrétisé en petits cubes, tout comme dans le cas du puits infini. La quantification de $\vec{k}$ entraîne *ipso facto* celle de l'énergie, donnée maintenant par l'expression suivante :

$$E_{n_x\, n_y\, n_z} \doteq \frac{\hbar^2 \vec{k}^2}{2\mu} = \frac{\hbar^2}{2\mu}\,(n_x^2 + n_y^2 + n_z^2)\frac{(2\pi)^2}{L^2} \qquad (n_u = 0, \pm 1, \pm 2, \pm 3, \ldots) \; . \tag{24.116}$$

Maintenant, le vecteur d'onde n'est plus restreint à l'octant positif. La comparaison des deux expressions (24.112) et (24.116) montre un facteur 4 d'écart, mais cette différence n'en induit aucune au niveau des résultats. En effet, dans le cas des sinus, le nombre d'états $n_{\mathrm{sinus}}(k)$ dont le vecteur d'onde a un module inférieur à k et contenus dans l'octant positif s'obtient divisant le volume de la sphère de rayon k par le volume d'un petit cube de côté π/L, soit $\delta^3 k = (\pi/L)^3$. On a donc :

$$n_{\mathrm{sinus}}(k) = \frac{1}{8}\,\frac{(4\pi/3)\,k^3}{(\pi/L)^3} = \frac{(kL)^3}{6\pi^2} \; . \tag{24.117}$$

[39]Si on prend $L = 1$ m, l'écart entre les premiers niveaux du puits infini (qui sont les moins denses) est de l'ordre de 10^{-20} eV !

[40]Le choix de la constante de normalisation sera fait d'ici peu.

Pour les ondes planes, il faut prendre la sphère entière, mais le volume élémentaire $\delta^3 k$ est maintenant $(2\pi/L)^3$ et on récupère le facteur 1/8. Au total, dans les sommations qui deviennent des intégrales, rien ne sera changé. Enfin, alors que la valeur $n_u = 0$ est exclue pour les sinus (la fonction d'onde serait identiquement nulle), rien ne l'interdit *a priori* pour les ondes planes. À la limite, l'adjonction d'un état singleton de mesure nulle ne modifie pas les résultats. En définitive, les sinus et les ondes planes produisent les mêmes prédictions physiques.

Par construction, l'espace $\vec{k}$ est alors discrétisé en petits cubes $\delta^3 k$ de volume $(2\pi/L)^3$, formant chacun ce que l'on peut appeler un *gros point*[41]. Ceci veut dire que toutes les sommations sur le vecteur $\vec{k}$ seront en temps utile remplacées par des intégrales, la somme discrète à L fini tendant vers une somme de Darboux ordinaire, point de départ servant à définir, par passage à la limite, l'intégrale de Riemann. En pratique :

$$\sum_{\vec{k}} (\ldots) \equiv \frac{1}{\delta^3 k} \sum_{\vec{k}} (\ldots)\, \delta^3 k \quad \rightarrow \quad \left(\frac{L}{2\pi}\right)^3 \int_{\mathbb{R}^3} (\ldots)\, \mathrm{d}^3 k \ . \tag{24.118}$$

Le facteur $[L/(2\pi)]^3$, inverse de $\delta^3 k$, est nécessaire ne serait-ce que pour des raisons d'homogénéité (les $(\ldots)$ dans les deux membres ont la même dimension, $\mathrm{d}^3 k$ a la dimension $(\text{longueur})^{-3}$) ; schématiquement, le rapport $\mathrm{d}^3 k/\delta^3 k$ tend vers 1 à la limite $L \rightarrow +\infty$. Bien évidemment, si tous les calculs sont faits convenablement, la longueur L doit disparaître tôt ou tard[42], pour ne produire finalement que des prévisions indépendantes des conditions aux limites précises, choisies uniquement en fonction de la commodité. Enfin, et c'est important, la densité d'états en énergie, $\rho(E)$, est la même, que l'on prenne des sinus comme en (24.113) ou des ondes planes. Ce point est essentiel puisque, en temps utile, on remplacera les sommations discrètes par des intégrales où apparaît ρ.

Cette discrétisation étant admise, il n'y a plus de difficulté technique. En particulier, chaque onde plane peut maintenant être normalisée sans problème :

$$\int_{\text{boîte}} |\psi_{\vec{k}}(\vec{r})|^2\, \mathrm{d}^3 r = 1 \quad \Longleftrightarrow \quad C = L^{-3/2} \quad \Longrightarrow \quad \psi_k(\vec{r}) = L^{-3/2}\, \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \ . \tag{24.119}$$

Au premier ordre, la probabilité de transition de $|\psi_1\rangle \rightarrow |\psi_{\vec{k}}\rangle$ est $P_{\vec{k}} = |A^{(1)}_{1\rightarrow\vec{k}}|^2$ où l'amplitude $A^{(1)}_{1\rightarrow\vec{k}}$ est donnée par (24.49), soit ici :

$$|A^{(1)}_{\mathrm{i}\rightarrow\vec{k}}| = \hbar^{-1} \left| \int_0^t \mathrm{d}t_1\, \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(E_{\vec{k}} - E_1)t_1} \langle \psi_{\vec{k}} | (w\, \mathrm{e}^{\mathrm{i}\omega t} + w^\dagger\, \mathrm{e}^{-\mathrm{i}\omega t}) | \psi_1 \rangle \right| \ . \tag{24.120}$$

Après intégration sur t_1 et élévation du module au carré, deux fonctions du type $f(T,\, \Omega)$ apparaissent (voir la définition (24.57)) ; la première vient du terme w et introduit $\Omega = \hbar^{-1}(E_{\vec{k}} - E_1) + \omega$, soit $f\big(t,\, \hbar^{-1}(E_{\vec{k}} - E_1) + \omega\big)$, la seconde vient de $w^\dagger$ et donne

[41] Ce type de pavage a été utilisé à propos du rayonnement thermique (voir Tome I, section 4.2).

[42] sauf, bien sûr, si on calcule des grandeurs extensives.

$f\big(t,\,\hbar^{-1}(E_{\vec{k}}-E_1)-\omega\big)$. Comme $\hbar\omega > 0$ et que $E_{\vec{k}} \gg E_1$, seul le terme contenant la deuxième fonction va contribuer en pratique au module carré de l'amplitude, d'où :

$$P_{\vec{k}} \simeq |A^{(1)}_{\mathrm{i}\to\vec{k}}|^2 = \hbar^{-2}\,|\langle\psi_{\vec{k}}|w^\dagger|\psi_1\rangle|^2\, f\big(t,\,\frac{E_{\vec{k}}-E_1}{\hbar}-\omega\big)\ . \tag{24.121}$$

Ceci est la probabilité élémentaire de trouver l'électron éjecté avec une impulsion $\vec{k}$ dont l'extrémité appartient au volume infiniment petit $\delta^3 k$. En pratique, le détecteur a une ouverture $\Delta^3 k$ assez petite pour avoir une bonne résolution, mais suffisamment grande pour produire un signal significatif et en tout cas certainement beaucoup plus grande que $\delta^3 k$ – qui peut être rendu petit *ad libitum.* Alors, le signal enregistré sur le détecteur est la somme ΔP des probabilités élémentaires associées aux événements satisfaisant $\|\vec{k}-\vec{k}_\mathrm{f}\| \le \Delta^3 k$, $\vec{k}_\mathrm{f}$ étant la valeur nominale de réglage du détecteur. P_f est :

$$P_\mathrm{f} = \sum_{\vec{k},\,\|\vec{k}-\vec{k}_\mathrm{f}\|\le\Delta^3 k} |A^{(1)}_{\mathrm{i}\to\vec{k}}|^2\ . \tag{24.122}$$

La somme sur $\vec{k}$ court par petits pas $\delta^3 k$ et peut être remplacée par une intégrale, conformément à (24.118). Il vient ainsi :

$$P_\mathrm{f} = \Big(\frac{L}{2\pi}\Big)^3 \int_{\|\vec{k}-\vec{k}_\mathrm{f}\|\le\Delta^3 k} \hbar^{-2}\,|\langle\psi_{\vec{k}}|w^\dagger|\psi_1\rangle|^2\, f\big(t,\,\frac{E_{\vec{k}}-E_1}{\hbar}-\omega\big)\,\mathrm{d}^3 k\ . \tag{24.123}$$

P_f est bien un nombre pur : f est homogène au carré d'un temps, l'élément de matrice est une énergie au carré. Selon les expressions (24.119) et (24.110), cet élément de matrice s'écrit précisément :

$$\langle\psi_{\vec{k}}|w^\dagger|\psi_1\rangle = \int_{\text{boîte}} L^{-3/2}\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\,\frac{\mathrm{i}e}{2\mu}A_0\,\vec{a}.\vec{p}\,\mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}}\,\psi_{100}(\vec{r}\,)\,\mathrm{d}^3 r\ , \tag{24.124}$$

soit, en remplaçant $\vec{p}$ par $-\mathrm{i}\hbar\vec{\nabla}$ et compte tenu du fait[43] que $\vec{a}.\vec{q}=0$:

$$\langle\psi_{\vec{k}}|w^\dagger|\psi_1\rangle = \frac{e\hbar}{2\mu}A_0\,L^{-3/2}\int_{\text{boîte}} \mathrm{e}^{+\mathrm{i}(\vec{q}-\vec{k}).\vec{r}}\,\vec{a}.\vec{\nabla}\,\psi_{100}(\vec{r}\,)\,\mathrm{d}^3 r\ . \tag{24.125}$$

En rapprochant mentalement (24.123) et (24.125), on voit qu'à cet endroit précis, tous les facteurs contenant L se compensent : on peut donc déjà anticiper la limite $L \to +\infty$ ici et là. Une intégration par parties (le terme tout intégré est nul) permet d'obtenir :

$$\langle\psi_{\vec{k}}|w^\dagger|\psi_1\rangle = \frac{e\hbar}{2\mathrm{i}\mu}A_0\,L^{-3/2}\,(2\pi\hbar)^{3/2}\,(\vec{q}-\vec{k}).\vec{a}\ \phi_{100}(\vec{p}=\hbar\vec{K})\ , \tag{24.126}$$

où $\phi_{100}(\vec{p})$ est la représentation-p de l'état 1s :

$$\phi_{100}(\vec{p}) = (2\pi\hbar)^{-3/2}\int_{\mathbb{R}^3} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\vec{p}.\vec{r}}\,\psi_{100}(\vec{r}\,)\,\mathrm{d}^3 r \equiv \chi_{10}(p)Y_{00}\ , \tag{24.127}$$

[43] ce qui exprime la commutation entre $\vec{p}$ et $\vec{A}$, aussi traduite par $\vec{\nabla}.\vec{A}=0$.

et où $\vec{K} = \vec{k} - \vec{q}$ est le transfert de moment du photon vers l'électron éjecté de l'atome. La fonction $\chi_{10}(p)$ se calcule facilement et a pour expression :

$$\chi_{10}(p) = \frac{16\pi}{(2\pi\hbar)^{3/2}} \frac{a_0^{3/2}}{[1 + (pa_0/\hbar)^2]^2} \ . \tag{24.128}$$

Avec tout ceci, (24.123) devient :

$$P_{\rm f} = \frac{8}{\pi}\Big(\frac{eA_0}{\mu}\vec{a}.\vec{K}\Big)^2 a_0^3 \int_{\|\vec{k}-\vec{k}_{\rm f}\| \leq \Delta^3 k} \frac{1}{(1 + a_0^2 K^2)^4} f(t, \omega_k - \omega_{\rm i} - \omega)\, k^2 {\rm d}k\, {\rm d}\Omega \ . \tag{24.129}$$

Compte tenu de $\omega_k = \frac{\hbar\vec{k}^2}{2\mu}$ et de $\omega_k \gg \omega_1 \equiv \hbar^{-1}|E_1|$, on a :

$$f(t, \omega_k - \omega_{\rm i} - \omega) \simeq f(t, \omega_k - \omega) \simeq 2\pi t\, \delta(\omega_k - \omega) = \frac{2\pi\mu}{\hbar k} t\, \delta\big(k - \sqrt{2\mu\omega/\hbar}\big) \ . \tag{24.130}$$

Au passage, on a supposé t suffisamment grand pour que $\delta\omega \equiv 1/t$ soit aussi petit que l'on veut, autorisant (comme on l'a fait pour établir la Règle d'or de Fermi) à remplacer le filtre approximatif f par une fonction δ convenablement ajustée (c'est-à-dire conformément à (24.58)). En reportant (24.130) dans (24.129), il vient :

$$P_{\rm f} = 16t\, \frac{a_0^3}{\mu\hbar} \big[eA_0\, \vec{a}.(\vec{k}_{\rm f} - \vec{q})\big]^2 \int_{\|\vec{k}-\vec{k}_{\rm f}\| \leq \Delta^3 k} \frac{k\, {\rm d}k\, {\rm d}\Omega}{[1 + a_0^2(k-q)^2]^4}\, \delta\big(k - \sqrt{2\mu\omega/\hbar}\big) \ . \tag{24.131}$$

Maintenant, de deux choses l'une :

- ou bien l'énergie du photon $\hbar\omega$ est différente de $\hbar^2 k_{\rm f}^2/(2\mu)$ – énergie nominale de la fenêtre du détecteur – et alors le domaine d'intégration de k est en-dehors du point de concentration de la fonction δ ; l'intégrale est alors nulle et $P_{\rm f} = 0$.
- ou bien l'énergie du photon est égale à la valeur affichée sur le détecteur et la fonction δ agit comme d'habitude.

Ceci traduit la conservation de l'énergie. Dans le second cas, la probabilité est :

$$P_{\rm f} = 16t\, \frac{a_0^3}{\mu\hbar} \big[eA_0\, \vec{a}.(\vec{k}_{\rm f} - \vec{q})\big]^2 \frac{k_{\rm f}}{[1 + a_0^2(k_{\rm f} - q)^2]^4}\, \Delta\Omega \ . \tag{24.132}$$

Il est agréable d'éliminer le potentiel-vecteur, afin ne faire apparaître que des quantités ayant un sens physique direct. L'énergie du champ dans la boîte de volume V est [195] :

$$E_{\rm champ} = \frac{\varepsilon_0}{2} \int_V \Big(\vec{\mathcal{E}}^2 + c^2 \vec{\mathcal{B}}^2\Big)\, {\rm d}^3 r = \varepsilon_0 \int_V \vec{\mathcal{E}}^2\, {\rm d}^3 r = N\hbar\omega \ , \tag{24.133}$$

où l'égalité $\mathcal{B}c = \mathcal{E}$ a été utilisée. La dernière égalité introduit N, nombre de photons d'énergie $\hbar\omega$ dans le (grand) volume $V = L^3$. Par ailleurs :

$$\vec{\mathcal{E}} = -\frac{\partial \vec{A}}{\partial t} = \omega A_0\, \vec{a} \cos(\vec{k}.\vec{r} - \omega t) \ , \qquad \overline{\vec{\mathcal{E}}}^2 = \frac{1}{2}\, \omega^2 A_0^2 \ , \tag{24.134}$$

où la barre horizontale désigne une moyenne dans le temps sur une période. La dernière égalité de (24.133) devient alors $\frac{1}{2}\varepsilon_0 V\omega^2 A_0^2 = N\hbar\omega$ et permet d'exprimer A_0^2 en fonction de la densité volumique de photons : N/V, $A_0^2 = \frac{2\hbar}{\varepsilon_0\,\omega}\,\frac{N}{V}$. En reportant ceci dans (24.132), il vient finalement :

$$\boxed{P_\mathrm{f} = 32\,t\,\frac{e^2\,a_0^3\,k_\mathrm{f}\,\left[\vec{a}.(\vec{k}_\mathrm{f} - \vec{q})\right]^2}{\varepsilon_0\mu\omega\,[1 + a_0^2(k_\mathrm{f} - q)^2]^4}\,\frac{N}{V}\,\Delta\Omega} \tag{24.135}$$

On note que le numérateur apporte une dépendance angulaire pour les photoélectrons. Il est d'ailleurs possible d'oublier $\vec{q}$, qui est très petit comparé à $\vec{k}_\mathrm{f}$; en effet :

$$\frac{\hbar^2 k_\mathrm{f}^2}{2\mu} = \hbar\omega = \hbar q c \quad \Longleftrightarrow \quad \frac{q}{k_\mathrm{f}} = \frac{\hbar k_\mathrm{f}}{2\mu c} = \frac{v_\mathrm{f}}{2c} \ll 1 \ , \tag{24.136}$$

puisque les électrons sont supposés non-relativistes. D'où la forme simplifiée :

$$P_\mathrm{f} \simeq 32\,t\,\frac{e^2\,a_0^3\,k_\mathrm{f}\,(\vec{a}.\vec{k}_\mathrm{f})^2}{\varepsilon_0\mu\omega\,(1 + a_0^2 k_\mathrm{f}^2)^4}\,\frac{N}{V}\,\Delta\Omega \ . \tag{24.137}$$

La probabilité de transition par unité de temps est simplement $p_\mathrm{f} = P_\mathrm{f}/t$. En utilisant le fait que :

$$(a_0 k_\mathrm{f})^2 = \left(\frac{\hbar^2}{\mu e'^2}\right)^2 \frac{2\mu E_\mathrm{f}}{\hbar^2} = \frac{E_\mathrm{f}}{|E_1|} \gg 1 \ , \qquad \hbar\omega \simeq E_\mathrm{f} \ , \tag{24.138}$$

on peut mettre p_f sous la forme approchée :

$$p_\mathrm{f} \simeq \frac{32}{\pi^2}\,\frac{c}{r_0}\lambda_\mathrm{C}^3\,\cos^2\theta\left(\frac{|E_1|}{E_\mathrm{f}}\right)^{5/2}\frac{N}{V}\Delta\Omega \ , \tag{24.139}$$

où r_0 et λ_C sont le rayon classique et la longueur d'onde Compton de l'électron :

$$r_0 = \frac{e'^2}{\mu c^2} \simeq 3\,\mathrm{F} \ , \qquad \lambda_\mathrm{C} = \frac{h}{\mu c} \simeq 0{,}024\ \text{Å} \ . \tag{24.140}$$

En définitive, (24.139) est l'expression de la probabilité de transition photoélectrique par unité de temps pour un $\vec{k}_\mathrm{f}$ à $\Delta\Omega$ près. θ est l'angle entre le vecteur d'onde final $\vec{k}_\mathrm{f}$ du photoélectron et la polarisation du champ électrique $\vec{E}$. Sans surprise, les électrons émis dans la direction du champ sont les plus nombreux : ce sont ceux qui vibrent le plus fortement sous l'effet de ce dernier ; quantitativement, le diagramme polaire de rayonnement est donné par $\cos^2\theta$ (aucun électron n'est émis perpendiculairement au champ électrique).

24.6 Le théorème de Gell-Mann et Low

Le théorème de Gell'Mann et Low [196] est fondamental : il établit une relation remarquable entre les états propres de deux Hamiltoniens, le second étant engendré par une montée infiniment lente (*adiabatique*) d'un opérateur perturbant le premier.

Plus précisément, les choses se présentent ainsi. Considérons le Hamiltonien dépendant du temps $H(t)$:

$$H(t) = H_0 + g\mathrm{e}^{\varepsilon t}\, V \qquad (\varepsilon > 0\,,\; t \leq 0)\ ; \tag{24.141}$$

g est à nouveau une constante de couplage sans dimension. Lorsque $t \to -\infty$, $H(t)$ tend vers H_0, et si $\varepsilon \to 0$, la montée est effectivement *infiniment* lente. L'idée est de calculer l'état à l'instant $t = 0$, quand l'état de départ est un certain état propre $|\Psi_0\rangle$ de H_0, d'énergie E_0 ; en d'autres termes, on va calculer $|\Psi(\varepsilon)\rangle$ défini comme :

$$|\Psi(\varepsilon)\rangle \stackrel{\text{déf}}{=} U_\varepsilon(0, -\infty)|\Psi_0\rangle\ ; \tag{24.142}$$

où $U_\varepsilon(t_1,\, t_2)$ est l'opérateur d'évolution en représentation - interaction ; introduisant le produit scalaire :

$$S(\varepsilon) \stackrel{\text{déf}}{=} \langle\Psi_0|\Psi(\varepsilon)\rangle \equiv \langle\Psi_0|U_\varepsilon(0, -\infty)|\Psi_0\rangle\ , \tag{24.143}$$

le théorème de Gell-Mann et Low affirme que si le rapport $\frac{1}{S(\varepsilon)}|\Psi(\varepsilon)\rangle$ existe à tous les ordres de la série de perturbation, sa limite $\varepsilon \to 0$:

$$\lim_{\varepsilon\to 0} \frac{U_\varepsilon(0, -\infty)|\Psi_0\rangle}{\langle\Psi_0|U_\varepsilon(0, -\infty)|\Psi_0\rangle} \equiv \lim_{\varepsilon\to 0} \frac{1}{S(\varepsilon)}|\Psi(\varepsilon)\rangle \stackrel{\text{déf}}{=} |\tilde{\Psi}\rangle\ , \tag{24.144}$$

est un état propre du Hamiltonien H :

$$\boxed{H|\tilde{\Psi}\rangle = \tilde{E}|\tilde{\Psi}\rangle} \tag{24.145}$$

En règle générale, ni le numérateur, ni le dénominateur n'ont une limite, seule celle du *rapport* existe, à condition qu'il ait un sens à tous les ordres de perturbation. Multipliant à gauche par $\langle\Psi_0|$ et puisque $H_0|\Psi_0\rangle = E_0|\Psi_0\rangle$, il vient :

$$\boxed{\tilde{E} = E_0 + \frac{1}{S}\langle\Psi_0|gV|\Psi\rangle} \tag{24.146}$$

La démonstration ci-dessous suit de près l'argument donné par Gell-Mann et Low [196] ; d'autres démonstrations sont disponibles (voir par exemple [197]).

L'opérateur $U_\varepsilon(t_2,\, t_1)$, noté U_1 antérieurement, satisfait ici (voir (24.20)) :

$$\mathrm{i}\hbar\,\frac{\partial}{\partial t}\,U_\varepsilon = V_\mathrm{I}(t)\,U_\varepsilon\ , \qquad V_\mathrm{I}(t) \stackrel{\text{déf}}{=} g\mathrm{e}^{\varepsilon t}\,\mathrm{e}^{-\frac{1}{\mathrm{i}\hbar}H_0 t}V\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}H_0 t}\ , \tag{24.147}$$

et a le développement[44] (voir (24.35)) :

$$U_\varepsilon(0, -\infty) = \sum_{n=0}^{+\infty} g^n \frac{(\mathrm{i}\hbar)^{-n}}{n!}\int_{-\infty}^0 \mathrm{d}t_1 \int_{-\infty}^0 \mathrm{d}t_2 \ldots \int_{-\infty}^0 \mathrm{d}t_n\, \mathrm{e}^{\varepsilon(t_1+t_2+\ldots+t_n)} \times \mathcal{T}\left[V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)\ldots\, V_\mathrm{I}(t_n)\right] \tag{24.148}$$

[44] Le terme $n = 0$ est l'identité $\mathbf{1}$.

Calculons maintenant plus précisément $(H_0 - E_0)|\Psi(\varepsilon)\rangle$, qui s'écrit :

$$(H_0 - E_0)|\Psi(\varepsilon)\rangle = (H_0 - E_0)U_\varepsilon(0, -\infty)|\Psi_0\rangle = [H_0,\, U_\varepsilon(0, -\infty)]|\Psi_0\rangle \ . \tag{24.149}$$

Soit le terme d'ordre n du développement (24.35), obtenu en y injectant le commutateur à droite dans (24.149), lequel introduit sous l'intégrale n-uple le commutateur suivant :

$$[H_0,\, V_\mathrm{I}(t_{i_1})V_\mathrm{I}(t_{i_2})...V_\mathrm{I}(t_{i_n})] = [H_0,\, V_\mathrm{I}(t_{i_1})]V_\mathrm{I}(t_{i_2})...V_\mathrm{I}(t_{i_n}) + V_\mathrm{I}(t_{i_1})[H_0,\, V_\mathrm{I}(t_{i_2})]V_\mathrm{I}(t_{i_3})...V_\mathrm{I}(t_{i_n}) + ... \ . \tag{24.150}$$

Par sa définition, l'opérateur V_I satisfait l'équation de Heisenberg relativement au Hamiltonien non perturbé H_0 (voir (24.23)), ce qui permet de remplacer chaque commutateur par une dérivée du genre $-\mathrm{i}\hbar\partial_{t_i}V_\mathrm{I}$. À ce stade, l'intégrale n-uple porte donc sur le terme :

$$-\frac{1}{(\mathrm{i}\hbar)^{n-1}}\frac{g^n}{n!}\sum_{\lambda=1}^{n!}P_\lambda\Big(\sum_{i=1}^{n}\frac{\partial}{\partial t_i}\Big)V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)...V_\mathrm{I}(t_n)\Theta(t_1-t_2)\Theta(t_2-t_3)...\Theta(t_{n-1}-t_n) \ , \tag{24.151}$$

où, conformément à (24.30), on a explicité l'action de l'opérateur chronologique. Il est aisé de voir que la dérivation du produit de fonctions Θ donne zéro, en raison de $(\partial_t + \partial_{t'})\Theta(t - t') = 0$. On peut donc extraire la somme des dérivées de la somme des permutations, pour la mettre en facteur, ce qui revient à faire commuter $\sum_i \partial_{t_i}$ avec l'opérateur chronologique. On arrive ainsi à :

$$-\frac{1}{(\mathrm{i}\hbar)^{n-1}}\frac{g^n}{n!}\int_{-\infty}^{0}\mathrm{d}t_1\int_{-\infty}^{0}\mathrm{d}t_2...\int_{-\infty}^{0}\mathrm{d}t_n\,\mathrm{e}^{\varepsilon(t_1+...+t_n)}\Big(\sum_{i=1}^{n}\frac{\partial}{\partial t_i}\Big)\mathcal{T}\big[V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)...V_\mathrm{I}(t_n)\big] \ . \tag{24.152}$$

En effectuant des permutations de variables muettes, on réalise que chaque dérivée donne la même contribution par intégration[45]. Il suffit donc de considérer un seul terme, ∂_{t_1} par exemple, puis de multiplier par n ; en effectuant une intégration par parties sur cette variable, on obtient deux termes ; le terme tout intégré est :

$$-\frac{1}{(\mathrm{i}\hbar)^{n-1}}\frac{g^n}{(n-1)!}\int_{-\infty}^{0}\mathrm{d}t_2...\int_{-\infty}^{0}\mathrm{d}t_n\,\mathrm{e}^{\varepsilon(t_2+...+t_n)}\mathcal{T}\big[V_\mathrm{I}(0)V_\mathrm{I}(t_2)...V_\mathrm{I}(t_n)\big] \ . \tag{24.153}$$

Le facteur $gV_\mathrm{I}(0)$ est égal à gV ; la sommation sur n de tous ces termes reconstitue $-gVU_\varepsilon(0,\,-\infty)|\Psi_0\rangle$, soit $-gV|\Psi(\varepsilon)\rangle$.

L'autre terme venant de l'intégration par parties est :

$$+\frac{\varepsilon}{(\mathrm{i}\hbar)^{n-1}}\frac{g^n}{(n-1)!}\int_{-\infty}^{0}\mathrm{d}t_1\int_{-\infty}^{0}\mathrm{d}t_2...\int_{-\infty}^{0}\mathrm{d}t_n\,\mathrm{e}^{\varepsilon(t_1+t_2...+t_n)}\mathcal{T}\big[V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)...V_\mathrm{I}(t_n)\big] \ ; \tag{24.154}$$

en écrivant $\frac{\varepsilon}{(\mathrm{i}\hbar)^{n-1}}\frac{g^n}{(n-1)!} = \mathrm{i}\hbar g\frac{\partial}{\partial g}\frac{\varepsilon}{(\mathrm{i}\hbar)^n}\frac{g^n}{n!}$, la sommation sur n de tous ces termes reconstitue à nouveau $|\Psi(\varepsilon)\rangle$. Rassemblant les deux termes venant de l'intégration par parties, il vient :

$$(H_0 - E_0)|\Psi(\varepsilon)\rangle = -gV|\Psi(\varepsilon)\rangle + \varepsilon\mathrm{i}\hbar g\frac{\partial}{\partial g}|\Psi(\varepsilon)\rangle \ , \tag{24.155}$$

[45]C'est à nouveau une conséquence du lemme de réarrangement pour le groupe symétrique $\mathcal{S}_n$.

soit :

$$(H - E_0)|\Psi(\varepsilon)\rangle = \varepsilon i\hbar g \frac{\partial}{\partial g}|\Psi(\varepsilon)\rangle \ . \tag{24.156}$$

En multipliant à gauche par $\frac{1}{S(\varepsilon)}\langle\Psi_0|$, on a :

$$\frac{1}{S(\varepsilon)}\langle\Psi_0|(H - E_0)|\Psi(\varepsilon)\rangle = \varepsilon i\hbar g \frac{1}{S(\varepsilon)}\langle\Psi_0|\frac{\partial}{\partial g}|\Psi(\varepsilon)\rangle = \varepsilon i\hbar g \frac{1}{S(\varepsilon)}\frac{\partial}{\partial g}S(\varepsilon) \ , \tag{24.157}$$

où a été utilisé le fait que $\langle\Psi_0|$ ne dépend pas de la constante de couplage g ; selon (24.146), le premier membre est[46] $\tilde{E} - E_0$, d'où l'égalité :

$$\tilde{E} - E_0 = \varepsilon i\hbar g \frac{1}{S(\varepsilon)}\frac{\partial S(\varepsilon)}{\partial g} \ . \tag{24.158}$$

Maintenant, multipliant (24.156) par $\frac{1}{S(\varepsilon)}$:

$$(H - E_0)\frac{1}{S(\varepsilon)}|\Psi(\varepsilon)\rangle = \varepsilon\frac{i\hbar g}{S(\varepsilon)}\frac{\partial}{\partial g}|\Psi(\varepsilon)\rangle = i\hbar\varepsilon g\frac{\partial}{\partial g}\Big(\frac{1}{S(\varepsilon)}|\Psi(\varepsilon)\rangle\Big) + \varepsilon\frac{i\hbar g}{S^2(\varepsilon)}\frac{\partial S(\varepsilon)}{\partial g}|\Psi(\varepsilon)\rangle \ . \tag{24.159}$$

soit :

$$\Big(H - E_0 - i\hbar\varepsilon g\frac{\partial}{\partial g}\Big)\frac{1}{S(\varepsilon)}|\Psi(\varepsilon)\rangle = \varepsilon\frac{i\hbar g}{S^2(\varepsilon)}\frac{\partial S(\varepsilon)}{\partial g}|\Psi(\varepsilon)\rangle = \frac{1}{S(\varepsilon)}(\tilde{E} - E_0)|\Psi(\varepsilon)\rangle \ , \tag{24.160}$$

où (24.158) a été utilisée pour écrire la dernière égalité ; finalement, on obtient :

$$\big(H - \tilde{E}\big)\frac{1}{S(\varepsilon)}|\Psi(\varepsilon)\rangle = i\hbar\varepsilon g\frac{\partial}{\partial g}\Big(\frac{1}{S(\varepsilon)}|\Psi\rangle(\varepsilon)\Big) \ . \tag{24.161}$$

On peut maintenant prendre la limite $\varepsilon \to 0$; le premier membre donne, par définition $(H - \tilde{E})|\tilde{\Psi}\rangle$ (voir (24.144)), limite supposée exister. Quant au second[47], sa limite doit aussi exister (puisque c'est le cas de celle du premier membre) ; comme il est multiplié par ε, cette limite est nulle, d'où le résultat :

$$\boxed{(H - \tilde{E})|\tilde{\Psi}\rangle = 0} \tag{24.162}$$

Ainsi, le théorème de Gell-Mann et Low définit un mode de construction d'un état propre du Hamiltonien perturbé ; cet état propre n'est pas forcément l'état fondamental de H, même si $|\Psi_0\rangle$ est celui de H_0. Indépendamment d'une mise en œuvre le plus souvent difficile[48], ce théorème n'est pas toujours applicable, puisqu'il suppose fondamentalement que la perturbation est *régulière*, au sens où les grandeurs d'intérêt sont des fonctions analytiques du paramètre de couplage g dans un voisinage de $g = 0$. Il existe des systèmes dont l'état fondamental *n'est pas* une fonction analytique de la constante de couplage.

[46] Sur l'égalité (24.158), on voit bien que la limite $\varepsilon \to 0$ n'est pas triviale : si on pouvait naïvement faire $\varepsilon = 0$ au second membre, on trouverait qu'il n'y a pas de déplacement d'énergie, ce qui serait visiblement stupide.

[47] L'existence de la limite du rapport $(\langle\Phi_0|\Psi(\varepsilon)\rangle^{-1}|\Psi(\varepsilon)\rangle$ signifie que, dans cette limite, le vecteur $|\Psi(\varepsilon)\rangle$ possède au moins une phase divergente.

[48] À titre d'exemple, voir [198].

24.7 Exercices et problèmes

24.7.1 Excitation coulombienne

Un projectile de masse M et de charge $Z_1 e$ est envoyé en direction d'un atome à Z électrons. Dans le cas où le projectile passe loin de l'atome, on peut traiter son mouvement de façon purement classique en le considérant comme une charge ponctuelle repérée par le vecteur dépendant du temps $\vec{R}(t)$; les électrons de la cible sont, eux, repérés par $\vec{r}_i$; m désigne la masse d'un électron. Initialement ($t = -\infty$), l'atome est dans son état fondamental d'énergie E_1.

1. Écrire $V(t)$, opérateur d'interaction entre la cible et le projectile.
2. Compte tenu de $r_i/R \ll 1\,\forall t$, écrire une expression simplifiée de $V(t)$; faire apparaître le moment dipolaire de l'atome, $\vec{d} = \sum_{i=1}^{Z} e\,\vec{r}_i$.
3. Soit b le paramètre d'impact du projectile, v_0 sa vitesse initiale ; son angle polaire ϕ varie de π (pour $t = -\infty$) à θ (pour $t = +\infty$). Utiliser la conservation du moment cinétique pour obtenir une relation entre R, ϕ, b et v_0.
4. En désignant par $|E_k\rangle$ les états propres de l'atome isolé, écrire au premier ordre l'expression de $A_k(t)$, amplitude de transition vers l'état $|E_k\rangle$ à l'instant t sous l'effet de la collision.
5. On se place désormais dans l'hypothèse d'une collision de durée très courte ; exprimer la condition correspondante en désignant par ξ la portée de l'interaction cible - projectile. En déduire une expression simplifiée de $A_k(+\infty)$.
6. En ne considérant que la diffusion vers l'avant ($\theta = 0$), trouver l'expression de la valeur moyenne de ΔE, énergie perdue par le projectile lors de la collision. En utilisant la relation (voir Tome I, problème 12.3.9) :

$$\sum_k |\langle E_k|\sum_{i=1}^{Z} u_i|E_1\rangle|^2\,(E_k - E_1) = Z\,\frac{\hbar^2}{2m} \qquad (u = x, y, z)\ , \tag{24.163}$$

montrer que la perte d'énergie ΔE est donnée par :

$$\Delta E = 2\,\frac{Z_1^2 Z e'^4}{m(bv_0)^2}\ . \tag{24.164}$$

24.7.2 Collision de deux spins 1/2

Deux spins 1/2 *discernables*, $\vec{S}_1$ et $\vec{S}_2$, interagissent suivant $V(t) = a(t)\,\vec{S}_1\vec{S}_2$; la fonction $a(t)$ est un scalaire et l'état initial ($t = -\infty$) des deux spins est $|m_1 = -1/2,\, m_2 = +1/2\rangle$.

1. (a) Montrer que le spin total est une constante du mouvement.
   (b) Quels sont les états finals possibles ?
   (c) Calculer au premier ordre de perturbation la probabilité $P_{\text{flip}}^{(1)}(t = +\infty)$ de basculement (*flip*) des spins sous l'effet de la collision (poser $a(t) = \hbar^{-1}\,\omega(t)$).
   (d) En introduisant les états propres du spin total, calculer exactement $P_{\text{flip}}^{(1)}(+\infty)$ et comparer à l'expression obtenue en 1c

2. Les deux spins sont maintenant de surcroît en interaction avec un champ magnétique $\mathcal{B}$ dirigé le long de Oz ; le Hamiltonien statique est $H_0 = -\mathcal{B}\,(\gamma_1 S_{1z} + \gamma_2 S_{2z})$.
   (a) Donner la nouvelle expression de $P_{\text{flip}}^{(1)}(+\infty)$.
   (b) En prenant $\omega(t) = \omega_0\,\mathrm{e}^{-(t/\tau)^2}$, calculer $P_{\text{flip}}^{(1)}(+\infty)$ explicitement et commenter sa variation en fonction de $\mathcal{B}$. A quelle condition l'approximation utilisée a-t-elle un sens ?

24.7.3 Impureté magnétique en phase solide

Un atome du genre alcalin (couches saturées et un électron périphérique s) se comporte essentiellement, à basse énergie, comme un spin $S = 1/2$; les états propres de $(\vec{S}^{\,2},\, S_z)$ sont notés $|\pm\rangle$. En présence d'un champ magnétique de module $\mathcal{B}$ parallèle à Oz, le Hamiltonien a la forme (σ_z est la matrice de Pauli usuelle) :

$$H_0 = \frac{1}{2}\hbar\omega_{\mathrm{L}}\sigma_z \qquad (\omega_{\mathrm{L}} > 0)\ . \tag{24.165}$$

1. (a) Donner l'expression de la pulsation de Larmor ω_{L} en fonction de la masse m de l'électron, de sa charge e, du facteur anormal g_{e} et de $\mathcal{B}$.
   (b) L'atome est soumis, entre 0 et t, à un champ de radio-fréquence ω ($t \gg \omega^{-1}$) ; la perturbation correspondante $V(t)$ est :

$$V(t) = v\,(\sigma_-\,\mathrm{e}^{+\mathrm{i}\omega t} + \sigma_+\,\mathrm{e}^{-\mathrm{i}\omega t}) \qquad (\omega > 0,\ v \in \mathbb{R})\ . \tag{24.166}$$

   Donner l'expression au premier ordre de P, probabilité de transition de l'état $|-\rangle$ vers l'état $|+\rangle$.
   (c) Le champ magnétique étant fixé, et l'atome étant dans l'état fondamental, on balaie en fréquence le champ excitateur. Illustrer par un schéma la variation de l'énergie absorbée en fonction de ω ? Quelle est la pulsation de résonance ?

2. Quand l'atome est inséré en phase solide, le spectre de résonance des sous-niveaux Zeeman est modifié comme indiqué sur la figure 24.4 ; on obtient maintenant deux raies voisines d'intensités différentes, dont il s'agit de comprendre l'origine.

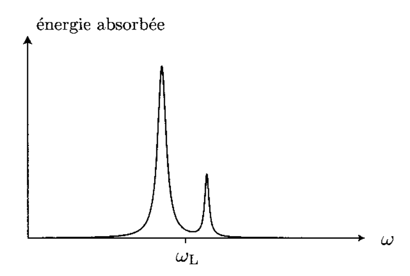


Figure 24.4: Variation de l'énergie absorbée en fonction de la fréquence du champ alternatif (phase solide)

(a) En phase solide, l'atome peut effectuer de petites vibrations autour de sa position d'équilibre, ce que l'on représente par un oscillateur harmonique à une dimension, de fréquence propre $\omega_0 > 0$. Dans ces conditions, le nouvel Hamiltonien est H_0' :

$$H_0' = \frac{1}{2}\hbar\omega_L\sigma_z + \hbar\omega_0\left(a^\dagger a + \frac{1}{2}\right) . \tag{24.167}$$

$a^\dagger$ et a sont les opérateurs de création et d'annihilation ; on désigne par $|\pm, n\rangle$ les états propres de H_0'. Faire le schéma des niveaux d'énergie de H_0' en distinguant les trois cas :

$$\omega_L \ll \omega_0 \ , \qquad \omega_L \sim \omega_0 \ , \qquad \omega_L \gg \omega_0 \ . \tag{24.168}$$

(b) Pour rendre compte de la structure du spectre en phase solide, on introduit un couplage entre les vibrations et le spin, représenté par l'opérateur W :

$$W = \hbar\Omega\,(a^\dagger\sigma_- + a\sigma_+) \qquad (\Omega > 0) \ . \tag{24.169}$$

Dans la suite, on se propose de diagonaliser exactement $H = H_0' + W$, en considérant exclusivement le sous-espace $\mathcal{E}$ engendré par les trois premiers niveaux de H_0', soit $|\psi_1\rangle = |-, 0\rangle$, $|\psi_2\rangle = |+, 0\rangle$ et $|\psi_3\rangle = |-, 1\rangle$. Représenter la variation en fonction de $\mathcal{B}$ des énergies de ces trois états.

(c) Calculer explicitement $W|\psi_i\rangle$ ($i = 1, 2, 3$) et en déduire que ce sous-espace est stable par H.

(d) $\mathcal{B}$ étant supposé différent de zéro, quel est l'état fondamental de H ?

(e) On écrit les deux autres vecteurs propres de H contenus dans $\mathcal{E}$ sous la forme :

$$|\psi_+\rangle = \cos\theta|+, 0\rangle + \sin\theta\,|-, 1\rangle \ , \quad |\psi_-\rangle = -\sin\theta\,|+, 0\rangle + \cos\theta\,|-, 1\rangle \ . \tag{24.170}$$

Les énergies propres associées aux $|\psi_\pm\rangle$ sont notées $E_\pm$. Donner les expressions de ces dernières (poser $\delta = (\omega_L - \omega_0)/2$) et tracer la variation des trois énergies propres de H en fonction de $\mathcal{B}$. Exprimer $\tan\theta$ à l'aide de Ω et de δ.

(f) Montrer que la perturbation définie par l'équation (24.166) peut induire maintenant *deux* transitions distinctes et donc, quand $\omega_L \simeq \omega_0$, expliquer qualitativement le spectre de la figure 24.4.

(g) Expliquer comment la constante de couplage Ω peut être déduite de la comparaison des intensités des deux raies.

24.7.4 Transitions à deux photons

Un atome est représenté par un système à trois niveaux[49] (voir fig. 24.5, $\omega_0 \neq \omega_0'$). À partir de $t = 0$, il est soumis à un champ variable $V(t) = v(t) + v^\dagger(t)$, où :

$$v(t) = \hbar e^{i\omega t} (\Omega|1\rangle\langle 2| + \Omega'|2\rangle\langle 3|) \qquad (\Omega,\, \Omega' \in \mathbb{R}) \ . \tag{24.171}$$

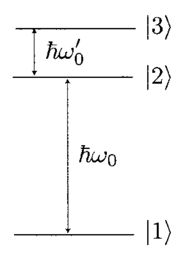


Figure 24.5: Système à trois niveaux

1. Calculer au premier ordre la probabilité de transition $P_{12}(t)$. Montrer que pour $t \gg |\omega_0 - \omega_0'|^{-1}$, $P_{12}(t)$ n'est non nulle que si $\omega \simeq \omega_0$ (résonance). Combien vaut à cet ordre la probabilité P_{13} ?
2. Calculer P_{13} au plus bas ordre non-nul.
3. En déduire l'allure schématique du spectre d'absorption quand l'atome est initialement dans l'état fondamental.

24.7.5 Retournement d'un spin par un champ magnétique éphémère

Un spin 1/2, de facteur gyromagnétique positif, est initialement dans l'état $|+1/2\rangle$, propre de S_z. Entre $t = 0$ et $t = T$, il est soumis à un champ magnétique de module $\mathcal{B}$ dirigé le long de Ox.

[49]Cette situation se rencontre pour l'atome d'hydrogène, dont le fondamental est un état ${}^2\mathrm{S}_{1/2}$. La dégénérescence des deux premiers états excités ${}^2\mathrm{P}_{1/2}$ et ${}^2\mathrm{S}^*_{1/2}$ est levée par les corrections radiatives (voir p. 898). La transition ${}^2\mathrm{S}^*_{1/2} \rightarrow {}^2\mathrm{S}_{1/2}$ est une transition à deux photons, ce qui explique la durée de vie très longue (de l'ordre de 1/7 e de seconde) de l'état ${}^2\mathrm{S}^*_{1/2}$. Au contraire, la transition ${}^2\mathrm{P}_{1/2} \rightarrow {}^2\mathrm{S}_{1/2}$ est la raie L_α de Lyman, la durée de vie de ${}^2\mathrm{P}_{1/2}$ étant de l'ordre de 10^{-9} s (voir aussi les commentaires p. 1074 à propos de l'effet Stark).

1. Trouver, au premier ordre, à $t > T$, la probabilité de transition vers l'état $|-1/2\rangle$.
2. Résoudre le même problème exactement ; préciser la condition permettant de se borner à un calcul au premier ordre.

24.7.6 Retour sur l'oscillation de Rabi

Soit un système à deux niveaux $|1\rangle$ et $|2\rangle$ séparés en énergie de $\hbar\omega_0$ (voir fig. 18.1) ; le Hamiltonien est :

$$H_0 = \frac{\hbar\omega_0}{2}\,(-|1\rangle\langle 1| + |2\rangle\langle 2|) \ . \qquad (24.172)$$

À partir de $t = 0$, où le système est dans l'état fondamental, on applique une perturbation $V(t)$ monochromatique donnée par :

$$V(t) = v(\mathrm{e}^{\mathrm{i}\omega t}\,|1\rangle\langle 2| + \mathrm{e}^{-\mathrm{i}\omega t}\,|2\rangle\langle 1|) \qquad (v = \hbar\Omega \in \mathbb{R}) \ . \qquad (24.173)$$

1. (a) Exprimer $H(t) = H_0 + V(t)$ à l'aide des matrices de Pauli.
   (b) Calculer la probabilité de transition $P_{12}(t)$ au premier puis au second ordre. Comment s'écrit la contribution d'ordre n à l'amplitude $A_{12}(t)$?
2. Comparer ces résultats approchés au traitement exact (voir problème 18.7.11).

24.7.7 Perturbations constante et gaussienne

On désigne par $U(t)$ l'opérateur d'évolution permettant de construire le vecteur d'état $|\Psi(t)\rangle$ à partir de l'état initial $|\Psi(0)\rangle$: $|\Psi(t)\rangle = U(t)|\Psi(0)\rangle$.

1. Quelle est l'équation satisfaite par $U(t)$? Que vaut $U(0)$?
2. Pour un Hamiltonien H indépendant du temps, quelle est l'expression de $U(t)$?
3. Avec $H = H_0 + V(t)$, H_0 étant indépendant du temps et $V(t)$ une perturbation non-nulle à $t > 0$, on pose $U(t) = U_0(t)U_1(t)$, où U_0 est l'opérateur d'évolution associé à H_0. Quelle est l'équation satisfaite par $U_1(t)$? Que vaut $U_1(0)$?
4. Expliquer comment on peut engendrer un développement de $U_1(t)$ en puissances de la perturbation $V(t)$.
5. Dans toute la suite, on arrête ce développement après l'ordre 1 en V. Donner l'expression approchée de U_1 qui en résulte.
6. Soit $|\psi_\mathrm{i}\rangle$ et $|\psi_\mathrm{f}\rangle$ deux états propres quelconques de H_0, d'énergies E_i et E_f respectivement. Donner l'expression au premier ordre de $A_{\mathrm{i}\to\mathrm{f}}$, amplitude de transition de l'état $|\psi_\mathrm{i}\rangle$ à l'état $|\psi_\mathrm{f}\rangle$ sous l'effet de la perturbation $V(t)$.

7. On considère les transitions d'un état $|\psi_\mathrm{i}\rangle$ donné vers des états propres $|\psi_m\rangle \neq |\psi_\mathrm{i}\rangle$ de H_0, d'énergies E_m. Les éléments de matrice de $V(t)$ sont de la forme :

$$\langle E_m|V(t)|\psi_\mathrm{i}\rangle = \theta(t)f(t)\,V_m \ , \tag{24.174}$$

où $\theta(t)$ est la fonction échelon-unité ($\theta(t<0)=0$, $\theta(t>0)=1$). Montrer que la probabilité de transition $P_{\mathrm{i}\to m}$ est de la forme $F(t)\,|V_m|^2$, où $F(t)$ est une certaine fonction à préciser

8. On pose $(E_m - E_\mathrm{i}) = \hbar\omega_{m\mathrm{i}}$. Déduire de ce qui précède la fonction $F(t)$ dans les deux cas suivants :

   (a) $f(t) = \mathrm{C^{ste}} = f_0$. Pour une durée T donnée de la perturbation, quelle est la largeur en pulsation de la bande d'états accessibles sous l'effet $V(t)$?

   (b) $f(t) = f_0\,\mathrm{e}^{-\frac{(t-t_0)^2}{2\tau^2}}$. Donner l'expression approchée de $F(t)$ quand $t \gg t_0 \gg \tau$. Quelle est l'énergie des états accessibles par $V(t)$?

24.7.8 Perturbation d'un oscillateur isotrope

Une particule de masse m et de charge q est liée harmoniquement dans $\mathbb{R}^3$, sa pulsation propre étant ω. Au départ ($t=-\infty$), la particule est dans son état fondamental.

On applique un champ électrique $\vec{\mathcal{E}}(t)$ dirigé suivant un axe appelé Oz, dont le module varie comme suit :

$$\mathcal{E}(t) = \mathcal{E}_0\mathrm{e}^{-(t/\tau)^2} \ . \tag{24.175}$$

Trouver, au premier ordre, la probabilité de transition vers un état excité quelconque.

24.7.9 Oscillateur en champ alternatif : traitements approché et exact

Un oscillateur harmonique linéaire, (masse m, charge q, pulsation propre ω_0) est soumis à un petit[50] champ électromagnétique dérivant du potentiel vecteur $\vec{A}$, dont la seule composante non nulle (écrite d'emblée dans l'approximation dipolaire électrique) est :

$$A(t) = A_0 \cos\omega t \ . \tag{24.176}$$

Le Hamiltonien H est :

$$H = \frac{1}{2m}[p - qA(t)]^2 + \frac{1}{2}m\omega_0^2x^2 \ . \tag{24.177}$$

Il s'agit de montrer que, moyennant des transformations unitaires judicieusement choisies, il est possible de se ramener à un problème *indépendant* du temps.

[50] La signification de ce qualificatif sera précisée ultérieurement. On rappelle que le champ électrique $\vec{\mathcal{E}}$ se déduit du potentiel vecteur $\vec{A}$ par $\vec{\mathcal{E}} = -\frac{\partial\vec{A}}{\partial t}$.

1. (a) On pose $x_0 = \frac{qE_0}{m\omega_0^2}$ avec $E_0 = \omega A_0$. Quelle est la dimension de x_0 ? Quelle est sa signification physique ?
   (b) Fabriquer une longueur, notée Δx, avec les trois grandeurs $\hbar$, m et ω_0.
   (c) Utiliser ces derniers résultats pour énoncer la condition entre A_0, q, m, ω_0 et ω exprimant que l'on se trouve dans une situation de champ faible, à toute fréquence raisonnable.

2. (a) Expliquer l'origine des différents termes apparaissant dans H.
   (b) Soit $\Psi(x,\, t)$ une fonction d'onde. On pose :

$$\Psi(x,\, t) = \mathrm{e}^{\mathrm{i}\alpha(t)}\, U_p(a)\, U_x(k)\, \Phi(x,\, t) \equiv U(a,\, k,\, \alpha)\, \Phi(x,\, t) \ , \tag{24.178}$$

   où $U_p(a) \stackrel{\text{déf}}{=} \mathrm{e}^{(\mathrm{i}/\hbar)a(t)p}$ et $U_x(k) \stackrel{\text{déf}}{=} \mathrm{e}^{\mathrm{i}k(t)x}$; α, a et k sont des scalaires, fonctions de la seule variable temps. À quelles conditions les deux transformations U_x et U_p sont-elles unitaires ? Que faut-il de plus pour que U le soit aussi[51] ?
   (c) Montrer que $\Phi(x,\, t)$ satisfait une équation de la forme :

$$\mathrm{i}\hbar\, \frac{\partial \Phi}{\partial t} = \tilde{H}\, \Phi \ , \tag{24.179}$$

   où $\tilde{H}$ est un opérateur dont on donnera l'expression formelle à l'aide de U, $\partial_t U$ et H.
   (d) Démontrer les relations suivantes, en précisant les deux fonctions f et g qui y apparaissent :

$$U_x^\dagger\, p\, U_x = p + g(k) \qquad U_x^\dagger\, x\, U_x = x \ , \tag{24.180}$$

$$U_p^\dagger\, x\, U_p = x + f(a) \qquad U_p^\dagger\, p\, U_p = p \ . \tag{24.181}$$

   (e) En utilisant les résultats de la question 2d, expliciter le terme de $\tilde{H}$ qui contient la dérivée partielle de U.
   (f) De même, expliciter l'autre terme apparaissant dans $\tilde{H}$.
   (g) Écrire les équations que doivent satisfaire les scalaires a, k et α pour que tous les termes linéaires en x et en p de $\tilde{H}$ s'annulent identiquement, ainsi que le terme ne dépendant que du temps.
   (h) Les équations obtenues en 2g étant satisfaites, exprimer $\Phi(x,\, t)$ en fonction de $\Psi_n(x,\, t)$, état stationnaire normalisé d'un oscillateur harmonique ordinaire.
   (i) Soit $F(x)$ une fonction infiniment dérivable. Combien vaut $U_p(a)\, F(x)$?
   (j) En déduire que les solutions de l'équation de Schrödinger de l'oscillateur soumis au champ se déduisent toutes de fonctions de la forme :

$$\Psi(x,\, t) = \mathrm{e}^{\mathrm{i}\alpha(t)}\, \tilde{\Phi}(x,\, t) \ , \tag{24.182}$$

   où $\tilde{\Phi}(x,\, t)$ est une fonction s'exprimant simplement à l'aide de $\Psi_n(x,\, t)$.

[51] Toutes ces conditions sont supposées réalisées dans la suite.

(k) Trouver[52] $k(t)$, puis $a(t)$ (faire apparaître x_0 dans l'expression de a).

3. (a) Sans faire aucun calcul, mais en justifiant le résultat, donner la valeur moyenne de la coordonnée $\langle x\rangle_n(t)$ dans l'état du type (24.182) construit avec $\Psi_n(x,\, t)$.
   (b) Quelle est la valeur moyenne de l'impulsion, $\langle p\rangle_n(t)$, dans le même état ?
   (c) Donner l'expression simplifiée de $a(t)$ dans les deux cas $\omega \ll \omega_0$ et $\omega \gg \omega_0$. Décrire le mouvement de l'oscillateur dans ces cas extrêmes.
   (d) Que se passe-t-il lorsque $\omega \to \omega_0$?
4. Reprendre brièvement les questions précédentes en traitant le champ électrique comme une perturbation.

24.7.10 Deux spins en interaction

Soit deux spins 1/2 de facteurs gyromagnétiques γ_1 et γ_2 soumis à un champ magnétique statique $\mathcal{B}$ et en interaction mutuelle par $V(t) \stackrel{\text{déf}}{=} \hbar^{-1}\omega(t)\vec{S}_1.\vec{S}_2$. L'état initial des deux spins est $|-\frac{1}{2},+\frac{1}{2}\rangle$. On posera :

$$\omega_{\rm L} \stackrel{\text{déf}}{=} \hbar(\gamma_1 - \gamma_2)\mathcal{B} \ . \tag{24.183}$$

1. Expliquer pourquoi le système se promène exclusivement dans le sous-espace engendré par les deux vecteurs $|-\frac{1}{2},+\frac{1}{2}\rangle$, $|+\frac{1}{2},-\frac{1}{2}\rangle$
2. On pose $\Psi(t)\rangle = a(t)|+\frac{1}{2},-\frac{1}{2}\rangle + b(t)|-\frac{1}{2},+\frac{1}{2}\rangle$. Écrire les équations différentielles pour les coefficients $a(t)$ et $b(t)$.
3. On introduit les fonctions auxiliaires $A(t)$ et $B(t)$ telles que :

$$a(t) = \mathrm{e}^{\frac{\mathrm{i}}{4}\int_{-\infty}^{t}\omega(t')\,\mathrm{d}t'}A(t) \ , \qquad b(t) = \mathrm{e}^{\frac{\mathrm{i}}{4}\int_{-\infty}^{t}\omega(t')\,\mathrm{d}t'}B(t) \ . \tag{24.184}$$

   Écrire les équations satisfaites par $A(t)$ et $B(t)$, et en déduire l'équation différentielle pour la fonction $r(t) \stackrel{\text{déf}}{=} \frac{A(t)}{B(t)}$.
4. On choisit une perturbation *carrée* agissant pendant la durée T :

$$\omega(t) = g\omega_{\rm L}\theta(t+\frac{T}{2})\theta(t-\frac{T}{2}) \ . \tag{24.185}$$

   Comme le Hamiltonien est statique, on sait résoudre directement ce problème pour a et b ; une autre méthode consiste à résoudre l'équation pour la fonction $r(t)$: trouver $r(T)$.
5. En déduire la probabilité P pour que le système soit trouvé dans l'état $|+\frac{1}{2},-\frac{1}{2}\rangle$ après la perturbation.

[52] On ne cherchera pas à écrire la solution la plus générale des équations obtenues en 2g ; pour le problème posé, il suffit d'obtenir une solution particulière.

6. Dans un gaz de spins, on peut modéliser les collisions entre spins par un terme du genre $V(t)$, avec $\omega(t)$ comme en (24.185). La durée des collisions, T, est supposée aléatoire, distribuée suivant la loi normale :

$$P(T) = \frac{1}{\sqrt{2\pi}\Delta T}\mathrm{e}^{-\frac{(T-\bar{T})^2}{2\Delta T^2}} \ . \tag{24.186}$$

Trouver la probabilité moyenne P pour que deux spins se retournent par collision.

24.7.11 Réponse linéaire d'un système dans un état pur : susceptibilité

On s'intéresse à une observable Ω d'un système perturbé par un certain opérateur $V(t)$, appliqué à partir d'un certain instant t_0 où l'état du système est $|\Psi(t_0)\rangle$. La valeur moyenne de Ω en présence de $V(t)$ est notée $\langle\Omega\rangle_V(t)$.

1. En utilisant la représentation-interaction, montrer que l'écart

$$\delta\Omega(t) \overset{\text{déf}}{=} \langle\Omega\rangle_V(t) - \langle\Omega\rangle_0(t) \tag{24.187}$$

est donné par :

$$\delta\Omega(t) = \frac{1}{\mathrm{i}\hbar}\int_{t_0}^{t} \langle\Psi_0(t')|\big[\Omega_{\mathrm{I}}(t-t'),\, V(t')\big]|\Psi_0(t')\rangle\,\mathrm{d}t' + \mathcal{O}(V^2) \ , \tag{24.188}$$

où $|\Psi_0(t)\rangle$ serait l'état du système à l'instant t si la perturbation n'avait pas été appliquée.

2. Dans le cas où $V(t) = v(t)V$, $v(t)$ étant un scalaire, on écrit la *réponse* $\delta\Omega(t)$ sous la forme :

$$\delta\Omega(t) = \int_{t_0}^{+\infty} \chi_{\Omega V}(t-t')v(t')\,\mathrm{d}t' + \mathcal{O}(V^2) \ , \tag{24.189}$$

où $\chi_{\Omega V}(t)$ est par définition la *susceptibilité* pour Ω quand le système est perturbé par V. Comment se simplifie l'expression de $\chi_{\Omega V}(t)$ quand $|\Psi(t_0)\rangle$ est l'un des états propres du système ?

3. La transformée de Fourier de $\chi_{\Omega V}(t)$ est $\hat{\chi}_{\Omega V}(\omega) \overset{\text{déf}}{=} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}\omega t}\chi_{\Omega V}(t)\,\mathrm{d}t$. Pouvez-vous expliquer l'affirmation : *la fonction $\hat{\chi}_{\Omega V}(z)$ est analytique dans le demi-plan complexe supérieur $\Im z \geq 0$?*

24.7.12 Réponse linéaire d'un système dans un état mixte : susceptibilité, fonctions de corrélation et relaxation

Un système est décrit par le Hamiltonien H_0, d'états propres notés $|n\rangle$: $H_0|n\rangle = \hbar\omega_n|n\rangle$. Dans l'état d'équilibre canonique, l'opérateur densité est $\rho(\beta) = Z^{-1}\mathrm{e}^{-\beta H_0}$ où Z désigne

la trace $\mathrm{Tr}\,\mathrm{e}^{-\beta H_0}$. Dans ce qui suit, on étudie la réponse *linéaire* de ce système lorsque, situé dans l'état d'équilibre à $t = -\infty$, il est soumis à une petite perturbation de la forme $V(t) \stackrel{\text{déf}}{=} v(t)V$. $v(t)$ est une fonction scalaire, V un opérateur statique purement non-diagonal, $\langle n|V|n\rangle = 0$ quel que soit n.

On note $H(t) \stackrel{\text{déf}}{=} H_0 + V(t)$, $U_0(t) \stackrel{\text{déf}}{=} \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}H_0 t}$, $\omega_{nm} \stackrel{\text{déf}}{=} \omega_n - \omega_m$, $\rho_n \stackrel{\text{déf}}{=} \langle n|\rho|n\rangle$. Par ailleurs, le symbole $\langle ...\rangle_{\text{eq}}$ désigne une valeur moyenne calculée avec l'état d'équilibre défini par $\rho(\beta)$.

1. Soit Ω une observable indépendante du temps dans la représentation de Schrödinger ; en l'absence de perturbation, la valeur moyenne $\langle\Omega\rangle_{\text{eq}}$ de Ω est constante (le démontrer) ; en présence de la sonde $V(t)$, cette valeur moyenne évolue dans le temps et on pose : $\delta\Omega(t) \stackrel{\text{déf}}{=} \langle\Omega\rangle(t) - \langle\Omega\rangle_{\text{eq}}$.

   (a) En partant de l'équation de Liouville, calculer l'écart $\delta\Omega(t)$ au premier ordre en V, et l'exprimer à l'aide de la susceptibilité $\chi_{\Omega V}$ précisément définie comme :
$$\chi_{\Omega V}(t-t') \stackrel{\text{déf}}{=} \frac{\mathrm{i}}{\hbar}\theta(t-t')\langle[\Omega_{\mathrm{I}},\, V_{\mathrm{I}}(t')]\rangle_{\text{eq}} \;; \tag{24.190}$$
   $\theta(t)$ est la fonction échelon-unité, l'indice $_{\mathrm{I}}$ désigne la représentation interaction des opérateurs.

   (b) Montrer que $\chi_{\Omega V}$ ne dépend bien que de la différence des temps ; s'assurer que $\delta\Omega$ est une quantité réelle et que la réponse linéaire est nulle pour une observable commutant avec H_0.

   (c) On note $\Omega_{nm} = \langle n|\Omega|m\rangle$, $V_{nm} = \langle n|V|m\rangle$; montrer que la susceptibiliité s'écrit aussi :
$$\chi_{\Omega V}(t-t') = \frac{1}{2\pi}\int_C \mathrm{e}^{-\mathrm{i}zt}\, G_{\Omega V}(z)\,\mathrm{d}z \;, \tag{24.191}$$
   où C est une droite allant de $-\infty$ à $+\infty$ juste au-dessus de l'axe réel, et où $G_{\Omega V}(z)$ est la fonction :
$$G_{\Omega V}(z) \stackrel{\text{déf}}{=} \hbar^{-1}\sum_{n,m}(\rho_n - \rho_m)\Omega_{nm}V_{mn}\,\frac{1}{\omega_{mn} - z} \;. \tag{24.192}$$

2. Sous l'effet de la perturbation, le système transite entre ses états propres et échange de l'énergie avec l'extérieur, son énergie moyenne $E(t) \stackrel{\text{déf}}{=} \mathrm{Tr}(H\rho)$ variant au cours du temps.

   (a) Montrer que, au plus bas ordre :
$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{\mathrm{d}v}{\mathrm{d}t}\int_{-\infty}^{+\infty} \chi_{VV}(t-t')v(t')\,\mathrm{d}t' \;. \tag{24.193}$$

(b) On choisit $v(t) = v_0 \mathrm{e}^{-\gamma t} \cos \omega_0 t$ ($\gamma > 0$), et on mesure la variation de l'énergie du système pendant un intervalle de temps $\Delta t \gg \omega_0^{-1}$. Montrer que la puissance $\mathcal{P}(t)$ ainsi mesurée est, dans la limite $\gamma \to 0$, égale à :

$$\mathcal{P}(t) = \frac{1}{4\mathrm{i}}\, a_0^2 \omega_0 \big[\tilde{\chi}_{VV}(\omega_0) - \tilde{\chi}_{VV}(-\omega_0)\big] \ , \tag{24.194}$$

où $\tilde{\chi}_{VV}(\omega)$ est la transformée de Fourier $\tilde{\chi}_{VV} \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} \chi_{VV}(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t$.

(c) Étudier la symétrie de $\tilde{\chi}_{VV}(\omega)$ et en déduire une forme simplifiée de $\mathcal{P}(t)$.

(d) Établir une relation simple entre $\tilde{\chi}_{VV}(\omega)$ et $G_{VV}(z)$.

(e) Montrer que pour un système à l'équilibre, il y a toujours absorption nette d'énergie.

3. Soit la fonction de corrélation d'équilibre symétrisée :

$$C_{\Omega V}(t,\, t') \stackrel{\text{déf}}{=} \frac{1}{2}\big[\langle \Omega_{\mathrm{I}}(t) V_{\mathrm{I}}(t') + V_{\mathrm{I}}(t')\Omega_{\mathrm{I}}(t)\big]\rangle_{\mathrm{eq}} \ ; \tag{24.195}$$

montrer que sa transformée de Fourier est égale à $\hbar \coth \frac{\beta\hbar\omega}{2}\, \Im\big[\tilde{\chi}_{\Omega V}(\omega)\big]$ quand Ω et V admettent une représentation réelle ou quand $\Omega = V$.

4. La fonction $G_{\Omega V}$ introduite plus haut permet aussi de décrire la dynamique de relaxation à partir d'un état *faiblement* hors d'équilibre. Pour établir ceci, on choisit maintenant $v(t) = v_0 \theta(-t)\, \mathrm{e}^{\gamma t}$ ($\gamma > 0$).

(a) Montrer que, quel que soit t, l'écart à l'équilibre est donné par :

$$t > 0 \ : \ \delta\Omega(t) = \mathrm{i}\frac{v_0}{2\pi} \int_C \mathrm{e}^{-\mathrm{i}zt} \frac{1}{z - \mathrm{i}\gamma}\, G_{\Omega V}(z)\, \mathrm{d}z \ , \tag{24.196}$$

et en déduire que l'on a $\delta\Omega(t) = -v_0 \mathrm{e}^{\gamma t} G(\mathrm{i}\gamma)$ si $t < 0$, et pour $t > 0$:

$$\langle \Omega \rangle(t) - \langle \Omega \rangle(0) = \frac{v_0}{2\pi} \int_0^t \mathrm{d}t' \int_C \mathrm{e}^{-\mathrm{i}zt'} \frac{z}{z - \mathrm{i}\gamma}\, G_{\Omega V}(z)\, \mathrm{d}z \ . \tag{24.197}$$

(b) Examiner la limite $\gamma \to 0$, correspondant à une montée adiabatique de la perturbation. En particulier, exprimer la relaxation de l'observable Ω à l'aide de la fonction de Kubo définie comme :

$$K_{\Omega V}(\beta,\, t) \stackrel{\text{déf}}{=} \beta^{-1} \int_0^\beta \langle V_{\mathrm{I}}(-\mathrm{i}\hbar\beta')\Omega_{\mathrm{I}}(t)\rangle_{\mathrm{eq}}\, \mathrm{d}\beta' \ . \tag{24.198}$$

5. *Application :* soit un gaz parfait de spins discernables $S_i = 1/2$. En présence d'un champ magnétique statique $\vec{\mathcal{B}}$ parallèle à Oz, le Hamiltonien est $H_0 = -\omega_{\mathrm{L}} \sum_i S_{zi}$, où ω_{L} est pulsation de Larmor, qui est proportionnelle à $\mathcal{B}$. La perturbation est un petit champ magnétique transverse $\vec{b}(t)$ polarisé suivant Ox.

(a) Calculer χ_{ux} ($u = x,\, y,\, z$), réponse linéaire associée à la valeur moyenne de S_u.

(b) Trouver la puissance $\mathcal{P}(t)$ pour un petit champ oscillant à la pulsation ω_0, et expliquer sa variation en température.

(c) Donner la dynamique à $t > 0$ de la petite composante transverse créée à $t = 0$. Examiner le cas limite de la montée adiabatique.

6. Commenter les variations des fonctions de corrélation $C_{ux}(t)$.

7. Déterminer la fonction de Kubo, et examiner la limite classique.

physique

Chapitre 25

Introduction à la description purement quantique de l'interaction champ-matière

Le but de ce chapitre est de montrer, par des moyens aussi élémentaires que possible, comment s'opère la quantification du champ électromagnétique. Trois applications importantes sont ensuite traitées : d'abord l'émission spontanée d'un atome, puis le passage graduel d'une oscillation à la Rabi, à une évolution irréversible en temps. Enfin, on montre comment l'interaction entre deux objets neutres (interaction de van der Waals) peut recevoir une interprétation électrodynamique (échange de photons virtuels).

Jusqu'à présent, le champ électromagnétique a été considéré comme un champ purement classique agissant de l'extérieur sur un système quantique (atome, par exemple). Cette description semi-classique est suffisante pour certaines applications mais est incapable de rendre compte de phénomènes importants, comme par exemple l'émission spontanée[1]. Un atome, un noyau, tout système porté dans un état excité n'y reste pas indéfiniment : même si le champ qui a excité le système a disparu entre temps, le système redescend au bout d'un certain temps à l'état fondamental, éventuellement en plusieurs étapes,

[1]La description semi-classique rend évidemment bien compte de l'émission *induite*. Dans le domaine optique, l'émission induite est en général beaucoup plus lente que l'émission spontanée et est souvent négligée de ce fait ; la situation est inversée dans la gamme des microondes – voir le commentaire juste après (21.78).

c'est–à-dire en transitant par d'autres niveaux intermédiaires. Ce temps est foncièrement aléatoire ; sa moyenne[2] est conventionnellement appelée durée de vie de l'état excité.

D'ailleurs, sur le plan des principes, on se doute qu'il va être tôt ou tard nécessaire de quantifier le champ électromagnétique, afin de construire un cadre théorique où champ et matière sont traités sur un même pied d'égalité. Il s'agit même d'un juste retour des choses : après tout, le rayonnement a joué un rôle crucial pour la génèse de la Théorie quantique, qu'il s'agisse du rayonnement du corps noir, de l'effet photoélectrique, de l'effet Compton, des spectres atomiques, etc. Il ne s'agit pas d'ailleurs seulement d'un juste retour des choses : la matière est constituée de charges en mouvement confiné ; quelle que soit la nature de la description précise de ces mouvements *accélérés*, on sait bien qu'en pareille situation il y a du rayonnement – l'un des problèmes[3] majeurs de la Physique classique (voir Tome I, chapitre 1, section 1.5 et problème 1.6.9). Il apparaît ainsi après coup que la formalisation élémentaire de la théorie quantique où le rayonnement reste en marge de la quantification est fondamentalement *incomplète*. L'inclusion quantique à part entière du champ électromagnétique fait d'ailleurs ressortir des aspects importants éclairant l'approche première. Notamment, on verra comment le choix d'une jauge (celle de Coulomb) fait émerger tout naturellement l'interaction électrostatique (de Coulomb), mise à la main en oubliant le champ dans les traitements moins élaborés. Dans cette jauge, en présence de charges, le champ n'est pas purement transverse ; sa partie longitudinale ne dépend que du potentiel scalaire U, tandis que toutes les composantes transverses sont uniquement fixées par le potentiel vecteur $\vec{A}$, transverse par définition de cette jauge.

La quantification du champ électromagnétique est une entreprise fort complexe dont l'achèvement rigoureux exige la mise en place d'un formalisme lourd et abstrait, que l'on peut trouver exposé dans tous ses détails dans l'ouvrage de Cohen-Tannoudji et al. [199], auquel le lecteur est renvoyé pour en savoir beaucoup plus. Il s'agit ici de présenter quelques unes parmi les idées importantes qui permettent de construire la théorie du système couplé photons - matière, appelée *Électrodynamique quantique*, suivant délibérément pour cela une argumentation pragmatique dont l'absence de rigueur formelle sera manifeste. On s'efforcera toutefois de ne sauter aucune étape importante, et de fournir des arguments plausibles quand la déduction formelle est hors de propos compte tenu des développements ou des techniques qu'elle exige.

Le système quantique ainsi défini, supposé isolé, sera décrit par un Hamiltonien arborant tout naturellement la forme :

$$H = H_{\text{at}} + H_{\text{champ}} + V \equiv H_0 + V \quad . \tag{25.1}$$

H_{at} est le Hamiltonien pris en compte pour la description semi-classique, alors noté H_0, H_{champ} est le Hamiltonien du champ libre ; quant à V, c'est le terme qui décrit

[2] On suppose qu'elle existe. Évidemment, il existe des variables aléatoires distribuées suivant des lois larges et qui sont dépourvues de moyenne (et *a fortiori* d'écart-type).

[3] À la réflexion, on peut se demander comment la Théorie quantique résout l'épineux problème de l'instabilité électrodynamique. Une réponse élémentaire : l'argument classique fait référence à la notion d'*accélération* d'une particule chargée, donc implicitement à la trajectoire réputée exister. En Mécanique quantique, il n'y a plus de trajectoire, seuls existent des états propres stables par construction, tant que le champ est laissé au dehors de la quantification.

l'interaction entre les deux parties du système. Par construction, l'atome et le champ forment un système isolé : H est indépendant du temps. De fait, l'analyse suivante fera apparaître des expressions réminiscentes de la théorie des perturbations stationnaires (par exemple (25.98)). Toutefois, l'objectif final est d'aller au-delà d'un schéma perturbatif afin d'obtenir, notamment, la décroissance exponentielle de la population d'un niveau atomique excité. Afin de montrer à la fois la puissance du formalisme et son universalité, on s'intéressera aussi à une interaction bien connue, celle de van der Waals, qui prévaut entre deux objets électriquement neutres (atomes, molécules) situés à grande distance l'un de l'autre par rapport à la dimension de chacun d'entre eux. On verra que cette interaction résulte d'un échange *virtuel* de photons [200], retrouvant ainsi l'image dynamique de processus virtuels déjà introduits dans la section 24.2.

25.1 Hamiltonien atome + champ

25.1.1 Lagrangien d'un champ scalaire classique unidimensionnel

A priori, un champ est un objet fort différent d'un système matériel formé, par exemple, d'un ensemble de particules ponctuelles pour lequel l'écriture du (d'un) Lagrangien ne pose aucune difficulté de principe. Si le système matériel est composé de N particules dans $\mathbb{R}^3$, son état dynamique est à chaque instant complètement déterminé par $3N$ coordonnées et par les vitesses correspondantes. Par contraste, un champ est une certaine fonction, scalaire, vectorielle,... d'*un* certain point de l'espace – là où le champ est calculé –, et du temps ; le champ le plus simple est ainsi une certaine fonction $u(x, t)$, dépendant d'*une* coordonnée x et du temps, t : il est scalaire (un seul *nombre* le caractérise), et *unidimensionnel*[4] puisqu'il n'y apparaît qu'une seule variable spatiale, x. Noter dès à présent que x n'est pas une variable *dynamique*, c'est juste ce qui permet de fixer le point de l'espace où le champ est calculé.

En fait, il existe des systèmes matériels qui, analysés dans une limite à préciser – traduisant physiquement une façon de les considérer à une certaine échelle –, conduisent tout naturellement à un champ. Comme le formalisme lagrangien est parfaitement clair et non-ambigu pour le système matériel discret, le passage à la limite permettra d'une part d'obtenir le champ associé, d'autre part le Lagrangien susceptible de décrire la dynamique de ce dernier[5].

L'exemple classique est celui d'une longue chaîne de boules et de ressorts[6] : $N \gg 1$ billes de masse m sur une ligne, chacune étant en interaction avec ses deux plus proches

[4]Par extrapolation, un champ $u(t)$ est souvent dit *zéro-dimensionnel*.

[5]Il est fréquent de considérer la procédure inverse : partant d'un champ continu, procéder à une discrétisation de l'espace pour écrire une théorie alors dite *sur réseau*. Une telle procédure est requise soit pour des raisons de commodité technique, soit, plus fondamentalement, quand la formulation d'emblée continue conduit à des ambiguïtés physiques. *Mutatis mutandis*, une théorie dynamique en temps continu impliquant des fonctions à mémoire courte modélisées par des fonctions $\delta(t - t')$ peut introduire subrepticement des difficultés liées à la violation en coulisse du Principe de causalité.

[6]situation que l'on retrouvera dans le chapitre 31 à propos des vibrations d'un solide et conduisant à la notion de *phonon*.

voisines. À l'équilibre, les billes sont immobiles et forment un réseau unidimensionnel parfait, la n^{e} bille étant au repos à l'abscisse $x_n = na$. S'agissant de décrire la dynamique des *petits* mouvements, le long de cette même ligne[7], autour de la configuration d'équilibre, on introduit l'écart u_n entre la position instantanée de la n^{e} bille et sa position d'équilibre, et la vitesse correspondante $v_n \stackrel{\text{déf}}{=} \dot{u}_n$; puis, on développe le potentiel d'interaction en se limitant au second ordre inclus par rapport aux écarts. Avec cette modélisation, on obtient une chaîne de boules reliées de proche en proche les unes aux autres par des ressorts parfaits ayant une certaine constante de raideur k – que l'on rapporte naturellement à la masse m en posant $k = m\omega^2$. Cela étant précisé, le Lagrangien s'en déduit immédiatement :

$$L(\{u_n\}_n, \{\dot{u}_n\}_n) = \sum_{n=1}^{N} \frac{1}{2} m\dot{u}_n^2 - \sum_{n=1}^{N} \frac{1}{2} m\omega^2 (u_{n+1} - u_n)^2 \ , \tag{25.2}$$

conduisant par variation de l'action aux équations de Lagrange ; une fois éliminées les vitesses $\dot{u}_n$, on obtient le système d'équations de Newton :

$$\boxed{\ddot{u}_n = -\omega^2 (2u_n - u_{n+1} - u_{n-1})} \tag{25.3}$$

les moments conjugués sont $p_n \stackrel{\text{déf}}{=} \frac{\partial L}{\partial \dot{u}_n} = m\dot{u}_n$, comme de coutume. Disposant de $2N$ conditions initiales u_{n0} et $\dot{u}_{n0}$, l'intégration de ce système donne une solution unique $u_n(t)$ décrivant complètement l'ensemble des déplacements à l'instant t résultant de cette configuration hors d'équilibre de départ.

L'idée à suivre maintenant consiste à regarder la chaîne à grande échelle (en s'en éloignant), de telle sorte que l'aspect discret du réseau de points s'efface peu à peu, ou en tout cas devient invisible avec une certaine *résolution* spatiale choisie d'avance : quand deux masses voisines sont trop proches au point de ne plus pouvoir les distinguer l'une de l'autre, autant les réunir dans un même "point". Formellement, ceci revient à réduire peu à peu le pas a du réseau d'équilibre et finalement à le faire tendre vers zéro. Pour que cette limite ait un sens physique, il faut déjà qu'un point ait une masse, ce qui revient à définir le rapport m/a comme une *constante* lors du passage à la limite ; pour la suite, on pose précisément $\mu \stackrel{\text{déf}}{=} m/a$. Dans ces conditions, le Lagrangien (25.2) prend la forme :

$$L(\{u_n\}, \{\dot{u}_n\}) = \sum_{n=1}^{N} \left[\frac{1}{2} \mu \dot{u}_n^2 - \frac{1}{2} \mu\omega^2 (u_{n+1} - u_n)^2 \right] a \ . \tag{25.4}$$

Introduisons maintenant une fonction $u(x)$ telle que :

$$u(x = na) \stackrel{\text{déf}}{=} u_n \ ; \tag{25.5}$$

ceci ne définit pas à proprement parler une *fonction*, sauf à admettre l'hypothèse suivant laquelle il existe de fait une fonction *lentement variable* à l'échelle du pas du réseau, auquel cas, la donnée de ses valeurs discrètes est, physiquement parlant, équivalente à

[7] Cette polarisation supposée des écarts n'est adoptée que pour la simplicité, et ne nuit pas à la généralité de l'argument suivant.

l'information contenue dans la fonction elle-même. La considération d'une telle fonction sera définitivement sans risque en adoptant enfin l'hypothèse suivant laquelle cette fonction est *continue* en x, ainsi que sa dérivée $\partial_x u$. Une condition est visiblement suffisante pour que ces hypothèses soient cohérentes : il suffit que la distribution des écarts initiaux, les u_{n0}, soit également *lentement* variable[8] à l'échelle a ; une telle propriété de l'état hors d'équilibre de départ, jointe à l'hypothèse des petites oscillations, doit convaincre que la variation lente le long du réseau est une propriété préservée au cours du temps.

Ces considérations autorisent à récrire le Lagrangien (25.4) comme suit :

$$L\big(u(x),\,\dot{u}(x)\big) = \sum_{n=1}^{N}\Big[\frac{1}{2}\mu\big(\dot{u}(x_n)\big)^2 - \frac{1}{2}\mu\omega^2[u(x_{n+1}) - u(x_n)]^2\Big]\,a\ . \tag{25.6}$$

Les hypothèses faites sur la fonction $u(x)$ permettent d'écrire :

$$u(x_{n+1}) - u(x_n) = a\Big[\frac{\partial u(x)}{\partial x}\Big]_{x=x_n} + a\varepsilon(a)\ , \qquad \lim_{a\to 0}\varepsilon(a) = 0\ ; \tag{25.7}$$

substituée dans le second terme de la sommation de (25.6), cette expression fait apparaître la quantité $(a\omega)^2$. Dans l'optique de la limite formelle $a \to 0$, il faut dire comment se comporte $a\omega$, homogène à une vitesse ; excluant évidemment que cette quantité diverge, on voit que la limite sera non-triviale ssi $a\omega$ a une valeur *constante* (ou en tout cas si sa limite est *finie*). Posant ainsi $v \stackrel{\text{déf}}{=} a\omega$, le Lagrangien (25.6) se récrit :

$$L(u,\,\dot{u}) = \sum_{n=1}^{N}\Big[\frac{1}{2}\mu\big(\dot{u}(x_n)\big)^2 - \frac{1}{2}\mu v^2\big(u'_x(x_n)\big)^2\Big]\,a\ , \tag{25.8}$$

où la notation simplifiée $u'_x \equiv \partial_x u$ a été introduite. Clairement, L apparaît maintenant comme une somme de Darboux constituant un précurseur de l'intégrale de Riemann. En passant maintenant à la limite $a \to 0$, $N \to +\infty$, on obtient :

$$\boxed{L[u,\,\dot{u}] = \int_{\mathbb{R}} \mathcal{L}[u,\,\dot{u}]\,\mathrm{d}x\ , \qquad \mathcal{L}[u,\,\dot{u}] \stackrel{\text{déf}}{=} \frac{1}{2}\mu\dot{u}^2 - \frac{1}{2}\mu v^2\,{u'_x}^2} \tag{25.9}$$

$\mathcal{L}$ étant, par définition, la *densité lagrangienne*. Le premier terme est de nature *cinétique* ; le second est un terme *potentiel* traduisant la raideur du champ : toute déformation par rapport à l'équilibre $u \equiv 0$ coûte de l'énergie. En définitive, partant d'un certain système matériel et prenant sa limite continue suivant une procédure contrôlée, on arrive à un champ $u(x)$ permettant de construire une densité lagrangienne. Une condition initiale étant donnée, la fonction $u(x,\,t)$ qui en descend est le champ décrivant, dans cette limite, les déplacements des points de la chaîne et obéissant à l'équation qui est la limite continue de (25.3) ; le second membre de celle-ci est $(v/a)^2[(u_{n+1} - u_n) - (u_n - u_{n-1})]$: le crochet

[8]Cette condition exclut ainsi une distribution du genre $u_{n0} = (-1)^n\delta$, correspondant à un réseau *dimérisé* ; la variation rapide de cette configuration fait pressentir qu'elle conduirait, dans la limite $a \to 0$, à une fonction *non-dérivable*. On reviendra dans le chapitre 31 sur ce type de question.

est l'accroissement des accroissements, c'est donc essentiellement une dérivée discrète *seconde* qui, munie du facteur a^{-2}, produit à la limite $\partial^2_{x^2}$:

$$\boxed{\frac{\partial^2 u}{\partial t^2} = v^2 \frac{\partial^2 u}{\partial x^2}} \tag{25.10}$$

Cette *équation de propagation* n'est finalement rien d'autre que *l'équation de Newton* du champ. Elle peut aussi s'obtenir par l'annulation de la variation de l'action S pour une variation arbitraire δu du champ u, en partant de la définition de l'action : $S \stackrel{\text{déf}}{=} \int_{t_\text{i}}^{t_\text{f}} \mathrm{d}t \int \mathcal{L}\,\mathrm{d}x$. La variation du terme cinétique est $\mu\dot{u}\delta\dot{u}$; une intégration par parties en temps (avec δu nul aux bornes) donne $-\ddot{u}\delta u$. En ce qui concerne le terme potentiel, il introduit la variation $-\frac{1}{2}\mu v^2 \delta \int {u'_x}^2\,\mathrm{d}x$; une intégration par parties en x donne $-\frac{1}{2}\mu v^2(-1)\int u''_{x^2}\delta u\,\mathrm{d}x$. Au total, on obtient :

$$\delta S = \int_{t_\text{i}}^{t_\text{f}} \mathrm{d}t \int (-\mu\ddot{u} + \mu v^2 u''_{x^2})\delta u(x)\,\mathrm{d}x \ ; \tag{25.11}$$

avec $-\mu\ddot{u} + \mu v^2 u''_{x^2} = 0$, la variation de l'action est bien nulle quel que soit l'incrément $\delta u(x)$: c'est l'équation de propagation ci-dessus. Le premier terme est $-\mu\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{u}}$; en *définissant* par ailleurs[9] :

$$\frac{\partial \mathcal{L}}{\partial u} \stackrel{\text{déf}}{=} \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial_x u} \tag{25.12}$$

on voit que l'équation dynamique prend la forme d'une équation de Lagrange :

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{u}} - \frac{\partial \mathcal{L}}{\partial u} = 0} \tag{25.13}$$

Poursuivant le programme habituel de la Mécanique analytique (voir Tome I, section 7.1), on définit ensuite le moment conjugué de $u(x)$, $\Pi(x)$, en recopiant au niveau de la densité lagrangienne la définition habituelle :

$$\boxed{\Pi(x) \stackrel{\text{déf}}{=} \frac{\partial \mathcal{L}}{\partial \dot{u}} = \mu\dot{u}(x)} \tag{25.14}$$

Quant à la densité hamiltonienne, on l'obtient par la transformation de Legendre :

$$\boxed{\mathcal{H} \stackrel{\text{déf}}{=} \Pi(x)\dot{u}(x) - \mathcal{L} = \frac{\Pi^2(x)}{2\mu} + \frac{1}{2}\mu v^2 {u'_x}^2} \tag{25.15}$$

Toutes ces expressions montrent bien que la variable x n'a rien de dynamique, mais est le simple substitut dans la version continue de l'indice n apparaissant dans la version discrète. Ce dernier, variant entre 1 et N avant la limite, représente le nombre de degrés

[9] L'égalité (25.12) doit être prise strictement comme une *définition*, permettant un raccourci d'écriture pour arriver à la forme (25.13). Le symbole $\frac{\partial \mathcal{L}}{\partial u}$ n'est pas la dérivée partielle de la densité lagrangienne par rapport à la variable u. Ces manipulations pourront sembler quelque peu formelles : elles peuvent être explicitées au prix de l'introduction de la notion de *dérivée fonctionnelle* [199].

de liberté de la chaîne finie ; c'est pourquoi on peut, en ce sens, dire qu'un champ (non zéro-dimensionnel !) est un système dont le nombre de degrés de liberté est égal à la puissance du continu. Pour ce champ continu, les variables dynamiques sont les deux grandeurs u et $\dot{u}$: ce sont elles qui qualifient complètement l'état du champ, et jouent les rôles d'une coordonnée et d'une vitesse[10].

S'agissant de la quantification du champ continu $u(x)$, elle s'effectue par passage à la limite de la version quantifiée du système discret. Pour obtenir celle-ci, il suffit d'associer aux grandeurs classiques u_n et $p_n = m\dot{u}_n$ des opérateurs satisfaisant :

$$[u_n,\, p_{n'}] = \mathrm{i}\hbar\delta_{nn'}\mathbf{1} \ , \tag{25.16}$$

soit, divisant par a membre à membre, $[u_n,\, \frac{1}{a}p_{n'}] = \mathrm{i}\hbar(\delta_{nn'}/a)\mathbf{1}$. Dans la limite continue, $n \to x$, $n' \to x'$, et $\delta_{nn'}/a \to \delta(x-x')$; la relation quantique fondamentale (25.16) se transforme ainsi en :

$$\boxed{\big[u(x),\, \Pi(x')\big] = \mathrm{i}\hbar\delta(x-x')\mathbf{1}} \tag{25.17}$$

25.1.2 Hamiltonien du champ libre

Revenant au champ électromagnétique, la première tâche à accomplir est de trouver son Hamiltonien (classique) H_{champ} ; l'étape de quantification ultérieure procédera comme à chaque fois que la quantification est en chantier : en s'inspirant de la description hamiltonienne classique et en posant des relations de commutation après avoir défini les opérateurs représentant les grandeurs physiques. On s'attend à voir émerger naturellement la notion de *photon*, grain de lumière se déplaçant à la vitesse c ; comme il n'existe pas de référentiel où le photon est au repos, ce fait permet de suspecter que le formalisme en fonction d'onde définie dans l'espace des coordonnées (représentation-q) risque d'être mal adapté – à supposer d'ailleurs qu'il existe[11]. En revanche, cela a un sens de parler d'un photon d'énergie ou d'impulsion bien définie. On pressent ainsi qu'il convient d'abord de formuler les équations classiques non dans l'espace réel mais dans l'espace *réciproque* des impulsions.

Les équations de Maxwell dans le vide s'écrivent :

$$\boxed{\vec{\nabla}.\vec{\mathcal{E}}(\vec{r},t) = 0 \qquad \vec{\nabla}.\vec{\mathcal{B}}(\vec{r},t) = 0} \tag{25.18}$$

$$\boxed{\vec{\nabla}\times\vec{\mathcal{E}}(\vec{r},t) = -\frac{\partial\vec{\mathcal{B}}}{\partial t} \qquad \vec{\nabla}\times\vec{\mathcal{B}}(\vec{r},t) = \frac{1}{c^2}\frac{\partial\vec{\mathcal{E}}}{\partial t}} \tag{25.19}$$

et doivent être considérées comme les équations du mouvement pour le champ. Introduisons maintenant les transformées de Fourier des champs :

$$\vec{\mathsf{E}}(\vec{k},\, t) \stackrel{\text{déf}}{=} \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}\vec{\mathcal{E}}(\vec{r},t)\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\,\mathrm{d}^3r \ , \quad \vec{\mathsf{B}}(\vec{k},\, t) \stackrel{\text{déf}}{=} \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}\vec{\mathcal{B}}(\vec{r},t)\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\,\mathrm{d}^3r \ ; \tag{25.20}$$

[10] Pour un point matériel classique, l'état dynamique est entièrement défini par la *trajectoire* ; réputée différentiable, elle est complètement déterminée à chaque instant par un point et la tangente ; il en va de même pour un champ classique.

[11] Il est en fait impossible de définir une fonction d'onde en représentation-q pour le photon.

les relations inverses sont :

$$\vec{\mathcal{E}}(\vec{r},t) = \int_{\mathbb{R}^3} \vec{\mathsf{E}}(\vec{k},\, t)\, \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\, \mathrm{d}^3k \ , \qquad \vec{\mathcal{B}}(\vec{r},t) = \int_{\mathbb{R}^3} \vec{\mathsf{B}}(\vec{k},\, t)\, \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\, \mathrm{d}^3k \ . \tag{25.21}$$

Comme les champs $\vec{\mathcal{E}}(\vec{r},t)$ et $\vec{\mathcal{B}}(\vec{r},t)$ sont réels, on a :

$$\vec{\mathsf{E}}^*(\vec{k},\, t) = \vec{\mathsf{E}}(-\vec{k},\, t) \ , \quad \vec{\mathsf{B}}^*(\vec{k},\, t) = \vec{\mathsf{B}}(-\vec{k},\, t) \ . \tag{25.22}$$

En appliquant les règles usuelles de dérivation avec les transformées de Fourier, les deux premières équations de Maxwell (25.18) se récrivent :

$$\vec{k}.\vec{\mathsf{E}}(\vec{k},\, t) = 0 \ , \qquad \vec{k}.\vec{\mathsf{B}}(\vec{k},\, t) = 0 \ , \tag{25.23}$$

où la transversalité du champ apparaît clairement par la nullité des produits scalaires avec le vecteur d'onde[12] $\vec{k}$. La transversalité autorisera à écrire en cas de besoin :

$$\vec{\mathsf{E}}(\vec{k},t) = \mathsf{E}_1(\vec{k},t)\vec{\varepsilon}_1(\vec{k}) + \mathsf{E}_2(\vec{k},t)\vec{\varepsilon}_2(\vec{k}) \ , \tag{25.26}$$

et de même pour la composante magnétique $\vec{\mathsf{B}}(\vec{k},t)$, les $\vec{\varepsilon}_i(\vec{k})$ étant deux vecteurs unitaires définissant avec $\vec{k}$ un repère trirectangle.

Quant à la paire d'équations (25.19), sa transcription dans l'espace de Fourier est :

$$\vec{k} \times \vec{\mathsf{E}}(\vec{k},\, t) = \mathrm{i}\frac{\partial \vec{\mathsf{B}}}{\partial t} \ , \qquad \vec{k} \times \vec{\mathsf{B}}(\vec{k},\, t) = \frac{1}{\mathrm{i}c^2}\frac{\partial \vec{\mathsf{E}}}{\partial t} \ . \tag{25.27}$$

Il se trouve que le couple $(\vec{\mathsf{E}}, \partial_t\vec{\mathsf{E}})$ (ou $(\vec{\mathsf{B}}, \partial_t\vec{\mathsf{B}})$) est plus naturel que $(\vec{\mathsf{E}}, \vec{\mathsf{B}})$; pour établir une équation fermée en $(\vec{\mathsf{E}}, \partial_t\vec{\mathsf{E}})$, on procède comme suit. En multipliant l'équation de droite de (25.27) vectoriellement par $\vec{k}$, et en utilisant la formule[13] du double produit vectoriel, on obtient :

$$\vec{\mathsf{B}}(\vec{k},\, t) = \frac{\mathrm{i}}{(kc)^2}\vec{k} \times \frac{\partial \vec{\mathsf{E}}}{\partial t} \ ; \tag{25.28}$$

---

[12] Le champ magnétique est toujours transverse ; en revanche, en présence d'une distribution de charge électrique $\rho(\vec{r},t)$, la première équation de Maxwell est :

$$\vec{\nabla}.\vec{\mathcal{E}}(\vec{r},t) = \frac{1}{\varepsilon_0}\rho(\vec{r},t) \ , \tag{25.24}$$

qui se transforme par Fourier en :

$$\vec{k}.\vec{\mathsf{E}}(\vec{k},\, t) = \frac{1}{\varepsilon_0}\tilde{\rho}(\vec{k},t) \ , \tag{25.25}$$

montrant que le champ électrique possède maintenant une composante longitudinale (i.e. le long de $\vec{k}$) non-nulle, que l'on notera $\vec{\mathcal{E}}_{\|}(\vec{r},\, t) \leftrightarrow \vec{\mathsf{E}}_{\|}(\vec{k},\, t)$.

Il est important de garder à l'esprit que la transversalité n'existe que si l'équation (25.18) est satisfaite *quel que soit* $\vec{r}$. En effet, il suffit qu'il y ait quelque part *une seule charge ponctuelle* pour que la composante $\vec{\mathcal{E}}_{\|}(\vec{r},\, t)$ soit non-nulle. À la réflexion, ceci n'est guère surprenant : l'existence d'une telle charge, à elle seule, brise la symétrie de translation dans l'espace, laquelle exclut, précisément, la possibilité d'une composante longitudinale.

[13] $\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a}.\vec{c})\vec{b} - (\vec{b}.\vec{a})\vec{c}$.

dérivant membre à membre en temps et reportant dans la première équation (25.27), on obtient :

$$\vec{k} \times \vec{\mathsf{E}}(\vec{k},\, t) = -\frac{1}{\omega^2}\vec{k} \times \frac{\partial^2 \vec{\mathsf{E}}}{\partial t^2} \ , \qquad \omega \stackrel{\text{déf}}{=} kc \ ; \tag{25.29}$$

cette égalité montre que les vecteurs $\vec{\mathsf{E}}$ et $\partial^2_{t^2}\vec{\mathsf{E}}$ diffèrent au plus par leurs composantes le long de $\vec{k}$, mais celles-ci sont nulles en raison de la transversalité (voir (25.23)). Il vient ainsi :

$$\boxed{\frac{\partial^2 \vec{\mathsf{E}}(\vec{k},\, t)}{\partial t^2} + \omega^2 \vec{\mathsf{E}}(\vec{k},\, t) = 0} \tag{25.30}$$

et de même pour la composante magnétique. Cette équation, remarquablement simple, évoque irrésistiblement celle d'un oscillateur harmonique : ceci ne doit pas constituer une surprise quand on se souvient que l'étude des modes normaux du champ libre dans une boîte (de taille quelconque) révèle que celui-ci constitue effectivement un ensemble d'oscillateurs harmoniques indépendants (voir Tome I, sous-section 4.2.2).

L'équation (25.30) est du second ordre en temps ; afin de faire un pas en direction du formalisme quantique[14], il est éclairant de faire apparaître une équation du premier ordre. Comme $\partial^2_{t^2} + \omega^2 = (\partial_t + \mathrm{i}\omega)(\partial_t - \mathrm{i}\omega)$, on voit que si on introduit un certain vecteur (complexe) $\vec{a}(\vec{k},\, t)$ proportionnel à $(\partial_t - \mathrm{i}\omega)\vec{\mathsf{E}}$, il satisfera l'équation du *premier* ordre[15] $(\partial_t + \mathrm{i}\omega)\vec{a} = 0$, qui a la forme voulue : $\mathrm{i}\partial_t \vec{a} = \omega\vec{a}$:

$$\boxed{\mathrm{i}\,\frac{\partial \vec{a}(\vec{k},\, t)}{\partial t} = \omega \vec{a}(\vec{k},\, t)} \tag{25.31}$$

Il ne manque plus qu'un facteur $\hbar$ des deux côtés pour que cette équation ait *formellement* l'allure d'une équation de... Schrödinger ou de Heisenberg[16]. La démarche précédente constitue une *"relecture quantique des équations de Maxwell"* [201]. L'équation de la variable $\vec{a}(\vec{k},\, t)$ est bien celle satisfaite par un oscillateur harmonique libre (i.e. non forcé) ; en effet, pour un tel oscillateur $(m,\, \omega)$, les équations de Hamilton sont $\dot{x} = \frac{p}{m}$ et $\dot{p} = -m\omega^2 x$, de sorte que la combinaison $a \propto x + \mathrm{i}\frac{p}{m\omega}$ satisfait précisément $\dot{a} + \mathrm{i}\omega a = 0$. Pour $\vec{k}$ fixé, $\vec{a}(\vec{k},\, t)$ est de fait un oscillateur harmonique : c'est une *variable normale* qui définit un *mode* du champ dont, selon Planck, le *quantum* indivisible vaut précisément $\hbar\omega$.

Cela fait, toutes les quantités (classiques) relatives au champ peuvent être exprimées en fonction du vecteur $\vec{a}(\vec{k},\, t)$. Posons précisément :

$$\vec{a}(\vec{k},\, t) = C(\omega)\Big(\frac{\partial \vec{\mathsf{E}}}{\partial t} - \mathrm{i}\omega\vec{\mathsf{E}}\Big) \ , \tag{25.32}$$

où $C(\omega)$ est une constante de proportionnalité à déterminer[17] ultérieurement (noter au

[14] où toutes les équations sont du *premier* ordre en temps.

[15] La procédure est réminiscente de celle suivie autour de l'équation de Dirac (chapitre 20).

[16] Acceptant l'assimilation formelle stricte, $\vec{a}(\vec{k},\, t)$ apparaît soit comme une *fonction d'onde* en représentation-p, soit comme un opérateur à identifier satisfaisant l'équation du mouvement de Heisenberg.

[17] C est *a priori* fonction de ω et de k, mais ces deux grandeurs sont reliées par $k = \omega/c$.

passage que $\vec{k}.\vec{a} = 0$). Inversement, on a :

$$\vec{\mathsf{E}}(\vec{k},\, t) = \frac{\mathrm{i}}{2C(\omega)}\left[\vec{a}(\vec{k},\, t) - \vec{a}^*(\vec{k},\, t)\right] \;, \quad \frac{\partial \vec{\mathsf{E}}(\vec{k},\, t)}{\partial t} = \frac{\omega}{2C(\omega)}\left[\vec{a}(\vec{k},\, t) + \vec{a}^*(\vec{k},\, t)\right] \quad (25.33)$$

Classiquement, l'expression de l'énergie du champ est [195] :

$$\mathcal{E}_{\text{class}} = \frac{\varepsilon_0}{2}\int_{\mathbb{R}^3} \left[\vec{\mathcal{E}}^2(\vec{r}, t) + c^2\vec{\mathcal{B}}^2(\vec{r}, t)\right] \mathrm{d}^3r \;; \quad (25.34)$$

La relation de Parseval - Plancherel est ici $\int_{\mathbb{R}^3} \vec{\mathcal{E}}^2(\vec{r}, t)\, \mathrm{d}^3r = (2\pi)^3 \int_{\mathbb{R}^3} \vec{\mathsf{E}}(\vec{k}, t)\vec{\mathsf{E}}^*(\vec{k}, t)\, \mathrm{d}^3k$, et de même pour la composante magnétique. Partant de (25.28), et utilisant la formule donnant le produit scalaire de deux produits vectoriels[18], on trouve :

$$\mathsf{B}(\vec{k}, t)\mathsf{B}^*(\vec{k}, t) = \frac{k^2}{\omega^4}\frac{\partial \vec{\mathsf{E}}}{\partial t}\frac{\partial \vec{\mathsf{E}}^*}{\partial t} \;; \quad (25.35)$$

reportant dans (25.34) :

$$\mathcal{E}_{\text{class}} = 4\pi^3\varepsilon_0 \int_{\mathbb{R}^3} \left[\vec{\mathsf{E}}(\vec{k}, t)\vec{\mathsf{E}}^*(\vec{k}, t) + \frac{k^2c^2}{\omega^4}\frac{\partial \vec{\mathsf{E}}(\vec{k}, t)}{\partial t}\frac{\partial \vec{\mathsf{E}}^*(\vec{k}, t)}{\partial t}\right] \mathrm{d}^3k \;; \quad (25.36)$$

exprimant maintenant le champ électrique en fonction de $\vec{a}$ (voir (25.33)), on obtient :

$$\mathcal{E}_{\text{class}} = 4\pi^3\varepsilon_0 \int_{\mathbb{R}^3} [C(\omega)]^{-2}\, \vec{a}(\vec{k},\, t).\vec{a}^*(\vec{k},\, t)\, \mathrm{d}^3k \;. \quad (25.37)$$

Il est à ce stade loisible de choisir le facteur de proportionnalité $C(\omega)$; il est naturel de vouloir faire disparaître le facteur $4\pi^3\varepsilon_0$, d'où l'étape $C^{-2} = X/(4\pi\varepsilon_0^3)$, où X est une grandeur à fixer, qu'il est naturel de prendre homogène à une énergie – auquel cas $\vec{a}$ sera homogène à la racine carrée d'un volume de $\mathbb{R}^3$. Dans la situation *préquantique* présente, le choix $X = \hbar\omega$ est tentant, et c'est celui que l'on retient pour finalement écrire :

$$\boxed{\mathcal{E}_{\text{class}} = \int_{\mathbb{R}^3} \hbar\omega\, \vec{a}(\vec{k},\, t).\vec{a}^*(\vec{k},\, t)\, \mathrm{d}^3k} \quad (25.38)$$

avec :

$$C(\omega) = \sqrt{\frac{4\pi^3\varepsilon_0}{\hbar\omega}} \;. \quad (25.39)$$

Cela étant, les composantes de Fourier des champs se récrivent (voir (25.28) et (25.33)) :

$$\vec{\mathsf{E}}(\vec{k},\, t) = \mathrm{i}\sqrt{\frac{\hbar\omega}{16\pi^3\varepsilon_0}}\left[\vec{a}(\vec{k},\, t) - \vec{a}^*(\vec{k},\, t)\right] \;, \quad \vec{\mathsf{B}}(\vec{k},\, t) = \mathrm{i}\sqrt{\frac{\hbar\omega}{16\pi^3\varepsilon_0 c^2}}\left[\vec{a}(\vec{k},\, t) + \vec{a}^*(\vec{k},\, t)\right] \quad (25.40)$$

Rappelons que toutes ces relations tiennent dans le vide, en l'absence de toute distribution de charge, situation dans laquelle le champ électrique est purement transverse.

[18] $(\vec{a} \times \vec{b}).(\vec{c} \times \vec{d}) = (\vec{a}.\vec{c})(\vec{b}.\vec{d}) - (\vec{a}.\vec{d})(\vec{b}.\vec{c})$.

La mise *à la main* de la constante de Planck peut paraître à ce stade très artificielle. On est en fait un peu dans la situation[19] de Planck : devoir introduire l'énergie d'un oscillateur harmonique, alors que la seule grandeur physiquement disponible est la pulsation ω de cet oscillateur. Cette nécessité se traduit par $E = \boxed{?} \times \omega$ où $\boxed{?}$ est homogène à une *action* ; en l'absence d'une action typique dans le problème, l'intrusion d'une constante fondamentale homogène à une action est tout compte fait inévitable – et ce ne peut être que $\hbar$.

Une fois exprimée l'énergie du champ classique, il est naturel de s'intéresser au vecteur de Poynting $\vec{\mathcal{S}}_{\text{class}}$, qui donne l'impulsion du champ. Partant de l'expression classique [195] :

$$\vec{\mathcal{S}}_{\text{class}} = \varepsilon_0 \int_{\mathbb{R}^3} \vec{\mathcal{E}}(\vec{r}, t) \times \vec{\mathcal{B}}(\vec{r}, t)\, \mathrm{d}^3 r \ , \tag{25.41}$$

des calculs analogues à ceux qui précèdent, complétés par le même choix (25.39) de la constante de proportionnalité, conduisent à :

$$\boxed{\vec{\mathcal{S}}_{\text{class}} = \int_{\mathbb{R}^3} \hbar \vec{k}\, \vec{a}(\vec{k}, t).\vec{a}^*(\vec{k}, t)\, \mathrm{d}^3 k} \tag{25.42}$$

Avoir écrit l'énergie du champ sous la forme (25.38) n'est pas, à proprement parler, avoir trouvé le Hamiltonien (classique). Celui-ci s'obtient en procédant comme d'habitude : former d'abord un Lagrangien conduisant à des équations du mouvement redonnant les équations de Maxwell ; puis, ayant défini les moments conjugués, construire le Hamiltonien par la transformation de Legendre habituelle.

On sait que les champs peuvent être exprimés (d'une infinité de façons) en fonction des potentiels (voir chapitre 17, section 17.4) $\vec{A}(\vec{r}, t)$ et $U(\vec{r}, t)$; pour mémoire, l'égalité entre $\vec{\mathcal{B}}$ et $\vec{A}$ est la suivante :

$$\vec{\mathcal{B}}(\vec{r}, t) = \vec{\nabla} \times \vec{A}(\vec{r}, t) \ ; \tag{25.43}$$

comme $\vec{\nabla} \times \vec{\nabla} = \vec{0}$, elle assure la satisfaction automatique de la deuxième équation de Maxwell (25.18). Posant ensuite :

$$\vec{\mathcal{E}}(\vec{r}, t) = -\frac{\partial \vec{A}(\vec{r}, t)}{\partial t} - \vec{\nabla} U(\vec{r}, t) \ , \tag{25.44}$$

c'est la troisième équation de Maxwell (25.19) qui est à son tour *ipso facto* satisfaite. Reportant ces expressions dans la quatrième équation de Maxwell, on obtient l'équation du mouvement pour le potentiel vecteur $\vec{A}(\vec{r}, t)$:

$$\frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} + \frac{1}{c^2}\vec{\nabla}\frac{\partial U}{\partial t} = \Delta \vec{A} - \vec{\nabla}(\vec{\nabla}.\vec{A}) \ . \tag{25.45}$$

[19]Autre situation analogue : la nécessité d'introduire le facteur h^{-DN} devant la fonction de partition d'un système *classique* à N degrés de liberté dans $\mathbb{R}^D$.

Compte tenu des relations (25.43) et (25.44), le champ est complètement déterminé si on connaît $\vec{A}(\vec{r},\, t)$ et $U(\vec{r},\, t)$. Comme l'équation pour $\vec{A}(\vec{r},\, t)$ est du second ordre en temps, la condition initiale requise pour avoir le potentiel vecteur est la donnée de $\vec{A}(\vec{r}, 0)$ et de $\partial_t \vec{A}(\vec{r}, 0)$, quel que soit $\vec{r}$. Quant au potentiel scalaire $U(\vec{r}, t)$, connaissant $\vec{A}(\vec{r}, t)$, on le trouve en utilisant l'équation :

$$\Delta U(\vec{r},\, t) = -\vec{\nabla}.\frac{\partial \vec{A}(\vec{r},\, t)}{\partial t} \ , \tag{25.46}$$

obtenue en reportant l'expression (25.44) du champ électrique dans la première équation de Maxwell (25.18). Le changement majeur entre les équations de Maxwell et celles donnant le mouvement des potentiels est le passage de l'ordre 1 à l'ordre 2 en temps : par référence à l'ordre des équations de Lagrange, ceci permet de comprendre que ce sont les potentiels qui, pour le champ électromagnétique, vont jouer le rôle des coordonnées. De fait, à l'instar de u et $\dot{u}$ pour le champ scalaire analysé dans la sous-section 25.1.1, les potentiels $U(\vec{r},\, t)$ et $\vec{A}(\vec{r},\, t)$ sont de fait les variables dynamiques qui déterminent complètement l'état du champ, leurs dérivées $\dot{U} \equiv \partial_t U$ et $\dot{\vec{A}} \equiv \partial_t \vec{A}$ faisant office de vitesses.

En définissant les transformés de Fourier $\vec{\mathsf{A}}(\vec{k},\, t)$ et $\mathsf{U}(\vec{k},\, t)$ des potentiels vecteur et scalaire, les relations entre champs et potentiels sont :

$$\vec{\mathsf{E}}(\vec{k},\, t) = -\frac{\partial \vec{\mathsf{A}}}{\partial t} - \mathrm{i}\vec{k}\mathsf{U}(\vec{k},\, t) \ , \qquad \vec{\mathsf{B}}(\vec{k},\, t) = \mathrm{i}\vec{k} \times \vec{\mathsf{A}}(\vec{k},\, t) \ ; \tag{25.47}$$

$\vec{\mathsf{A}}(\vec{k}, t)$ et $\mathsf{U}(\vec{k}, t)$ satisfont également les relations de conjugaison analogues à (25.22) : $\vec{\mathsf{A}}^*(\vec{k}, t) = \vec{\mathsf{A}}(-\vec{k}, t)$ et $\mathsf{U}^*(\vec{k}, t) = \mathsf{U}(-\vec{k}, t)$; il en résulte que si l'on veut, par symétrie, pouvoir considérer que le potentiel (scalaire et vecteur) et son complexe conjugué comme des grandeurs *indépendantes*, il faut restreindre toute éventuelle intégration sur le vecteur $\vec{k}$ à la *moitié* de $\mathbb{R}^3$. Par ailleurs, la seconde équation (25.47) montre que le champ magnétique ne voit que les composantes *transverses* du potentiel vecteur.

Le relatif arbitraire des potentiels vecteur et scalaire autorise à faire un choix inspiré par la simplicité technique ; la considération de (25.45) suggère ainsi fortement de prendre :

$$\boxed{\vec{\nabla}.\vec{A} = 0} \tag{25.48}$$

ceci définit la *jauge de Coulomb* qui, en outre, conduit tout naturellement à une interprétation physique particulièrement transparente, comme on le verra dans la suite[20]. Revenant à la première équation (25.47), on voit que, en jauge de Coulomb où $\vec{A}$ est transverse par définition, la seule contribution au champ électrique longitudinal vient du potentiel scalaire U. Autrement dit, pour le champ libre où tout est transverse, le choix de la jauge de Coulomb permet de prendre U identiquement nul. Dans cette jauge, et pour le champ transverse, le potentiel vecteur satisfait donc :

$$\frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} - \Delta \vec{A} = 0 \ ; \tag{25.49}$$

[20] En présence de charges, la composante transverse du champ est liée à $\vec{A}$, alors que la composante longitudinale ne dépend que de U.

avec le transformé de Fourier $\vec{\mathsf{A}}$, les deux équations précédentes se récrivent :

$$\boxed{\vec{k}.\vec{\mathsf{A}} = 0 \qquad \frac{\partial^2\vec{\mathsf{A}}}{\partial t^2} + \omega^2\vec{\mathsf{A}} = 0} \tag{25.50}$$

La question se pose maintenant de trouver le Lagrangien du champ, toujours supposé libre. Des arguments généraux [202] permettent de le construire *ab initio* ; on se contentera ici d'une démarche plus pragmatique, en posant une expression du Lagrangien et en montrant que les équations de Lagrange qui en résultent redonnent bien les équations de Maxwell. Soit donc le Lagrangien (appelé *Lagrangien standard* [199]) :

$$L_{\text{champ}} \stackrel{\text{déf}}{=} \frac{\varepsilon_0}{2}\int_{\mathbb{R}^3}\left[\vec{\mathcal{E}}^2(\vec{r}\,) - c^2\vec{\mathcal{B}}^2(\vec{r}\,)\right]\mathrm{d}^3r \stackrel{\text{déf}}{=} \int \mathcal{L}_{\text{champ}}\,\mathrm{d}^3r \ , \tag{25.51}$$

où l'égalité de droite définit la densité lagrangienne du champ libre. Il s'agit maintenant de montrer que l'équation d'Euler - Lagrange avec cette densité conduit à des équations de Lagrange équivalentes aux équations de Maxwell. Compte tenu de la simplicité de celles-ci quand on se place dans l'espace de Fourier, il est tout indiqué de définir une densité lagrangienne dans cet espace (pris seulement dans sa moitié, voir p. 1152), notée $\tilde{\mathcal{L}}$. Par la relation de Parseval - Plancherel, on a[21] :

$$\tilde{\mathcal{L}} = 2\times\frac{\varepsilon_0}{2}\left[|\vec{\mathsf{E}}(\vec{k}\,)|^2 - c^2|\vec{\mathsf{B}}(\vec{k}\,)|^2\right] \equiv \varepsilon_0\left[\vec{\mathsf{E}}(\vec{k}\,)\vec{\mathsf{E}}(-\vec{k}\,) - c^2\vec{\mathsf{B}}(\vec{k}\,)\vec{\mathsf{B}}(-\vec{k}\,)\right] \ , \tag{25.52}$$

En vertu de (25.47) avec $\mathsf{U}\equiv 0$, il vient $\tilde{\mathcal{L}} = \varepsilon_0\left[|\dot{\vec{\mathsf{A}}}(\vec{k}\,)|^2 - c^2|\vec{k}\times\vec{\mathsf{A}}(\vec{k}\,)|^2\right]$; compte tenu de la transversalité de $\vec{\mathsf{A}}$ et de l'égalité rappelée dans la note 18, il vient :

$$\tilde{\mathcal{L}} = \varepsilon_0\Big[\dot{\vec{\mathsf{A}}}(\vec{k}\,)\,\dot{\vec{\mathsf{A}}}^*(\vec{k}\,) - c^2[\vec{k}.\vec{\mathsf{A}}(\vec{k}\,)][\vec{k}.\vec{\mathsf{A}}^*(\vec{k}\,)]\Big] \ ; \tag{25.53}$$

Pour le champ, ce sont les potentiels qui jouent le rôle des coordonnées, leurs dérivées étant les vitesses. Pour le potentiel scalaire, il n'y a pas lieu d'écrire une telle équation : en jauge de Coulomb avec $U\equiv 0$, la première équation $\vec{k}.\vec{\mathsf{E}} = 0$ est automatiquement satisfaite, et d'ailleurs $\tilde{\mathcal{L}}$ ne dépend ni de U ni de $\dot{\mathsf{U}}$.

En désignant par $\mathsf{A}_\alpha(\vec{k}\,)$ l'une des composantes de $\vec{\mathsf{A}}(\vec{k}\,)$, l'équation de Lagrange relative au potentiel vecteur est :

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\tilde{\mathcal{L}}}{\partial\dot{\mathsf{A}}_\alpha^*} - \frac{\partial\tilde{\mathcal{L}}}{\partial\mathsf{A}_\alpha^*} = 0 \ , \tag{25.54}$$

Le moment conjugué de A_α^* est par définition $\Pi_{\mathsf{A}_\alpha^*} \stackrel{\text{déf}}{=} \frac{\partial\tilde{\mathcal{L}}}{\partial\dot{\mathsf{A}}_\alpha^*} = \varepsilon_0\dot{\mathsf{A}}_\alpha$, égal à $-\varepsilon_0\vec{\mathsf{E}}$ d'après (25.47) ; selon (25.53), $\frac{\partial\tilde{\mathcal{L}}}{\partial\mathsf{A}_\alpha^*}$ vaut $-\varepsilon_0c^2(\vec{k}.\vec{\mathsf{A}})\,k_\alpha$, d'où l'équation de Lagrange pour $\vec{\mathsf{A}}^*$: $\dot{\vec{\mathsf{E}}} = c^2(\vec{k}.\vec{\mathsf{A}})\,\vec{k} \equiv c^2\vec{k}\times\vec{\mathsf{B}}$: c'est la quatrième équation de Maxwell.

[21] Le facteur 2 vient de la relation de Parseval - Plancherel écrite avec $\int_{\frac{1}{2}\mathbb{R}^3}\mathrm{d}^3k$.

Une fois établi que le Lagrangien standard (25.51) est un candidat permettant de retrouver les équations dynamiques (celles de Maxwell), le Hamiltonien s'obtient par la transformation de Legendre habituelle en formant $\sum_\alpha \Pi_{\mathsf{A}_\alpha^*} \dot{\mathsf{A}}_\alpha^* - \mathcal{L} + \text{c.c.}$; le premier terme est $-\varepsilon_0 \vec{\mathsf{E}}(-\vec{\mathsf{E}}^*)$. Revenant à (25.52), puis en utilisant à nouveau la relation[22] de Parseval - Plancherel, on trouve finalement :

$$H_{\text{champ, cl}} \stackrel{\text{déf}}{=} \frac{\varepsilon_0}{2} \int_{\mathbb{R}^3} \left[\vec{\mathcal{E}}^2(\vec{r}) + c^2 \vec{\mathcal{B}}^2(\vec{r})\right] \mathrm{d}^3 r \tag{25.55}$$

Cette expression coïncide avec celle qui donne l'énergie classique du champ, éq. (25.34), et entraîne notamment que l'expression de $H_{\text{champ, cl}}$ à l'aide des variables normales $\vec{a}(\vec{k}, t)$ est donnée par le second membre de (25.38) ; cette identité valide le formalisme lagrangien et hamiltonien brièvement exposé ci-dessus, et peut ainsi servir de point de départ à la quantification du champ.

Comme très souvent, l'usage de la transformation de Fourier dans $\mathbb{R}^3$ peut conduire à des difficultés techniques, puisque le module du champ ne tend pas vers zéro à l'infini. La difficulté se contourne comme on l'a fait lors du traitement de l'effet photoélectrique (chapitre 24, sous-section 24.5.2) : on enferme le champ dans une grande boîte cubique de côté L ; comme des conditions aux limites franches introduisent des complications techniques, et que la physique ne doit pas ici dépendre de ce qui se passe précisément sur les parois de la boîte de confinement, on préfère utiliser des conditions aux limites périodiques[23] en identifiant $\mathrm{e}^{\mathrm{i}k_u(u+L)} = \mathrm{e}^{\mathrm{i}k_u}$, où $u = x, y, z$. Il en résulte $\mathrm{e}^{\mathrm{i}k_u L} = 1$, ce qui discrétise l'espace réciproque en petits grains cubiques de volume $(2\pi/L)^3$, et permet l'identification $\int \mathrm{d}^3 k \longleftrightarrow (2\pi/L)^3 \sum_{\vec{k}}$. Avec cette discrétisation et en introduisant les deux composantes $a_{\vec{k}\vec{\varepsilon}}$ de $(2\pi/L)^3\, \vec{a}(\vec{k})$ sur les vecteurs de polarisation comme en (25.26), le Hamiltonien classique du champ libre prend finalement la forme :

$$\boxed{H_{\text{class}} = \sum_{\vec{k}, \vec{\varepsilon}} \hbar\omega_k\, a_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}}^*} \tag{25.56}$$

avec[24] $\omega_k = kc$; noter que $a_{\vec{k}\vec{\varepsilon}}$ est maintenant sans dimension.

La quantification peut maintenant s'effectuer de façon pragmatique en revenant au sens des variables $a_{\vec{k}\vec{\varepsilon}}$ qui, tout comme $a_\varepsilon(\vec{k})$, sont associées à un mode d'oscillateur harmonique ; on a vu, par référence avec un oscillateur matériel (m, ω), que la variable normale $a_{\vec{k}\vec{\varepsilon}}$ est proportionnelle à $x + \mathrm{i}\frac{p}{m\omega}$, combinaison qui est elle-même proportionnelle à l'opérateur d'annihilation a de cet oscillateur. *Mutatis mutandis*, on est ainsi conduit d'une part à symétriser l'expression de H_{class} en remplaçant le produit par la somme $\frac{1}{2}(a_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}}^* + a_{\vec{k}\vec{\varepsilon}}^* a_{\vec{k}\vec{\varepsilon}})$, d'autre part à substituer à la variable normale $a_{\vec{k}\vec{\varepsilon}}$ et $a_{\vec{k}\vec{\varepsilon}}^*$

[22] L'écriture de Parseval - Plancherel avec $\int_{\mathbb{R}^3} \mathrm{d}^3 k$ réintroduit un facteur $\frac{1}{2}$.

[23] Aucune propriété physique ne devant dépendre de L, cette longueur artificielle doit disparaître de l'expression de toute grandeur physique en fin de calcul. En particulier, toutes les grandeurs extensives doivent contenir en facteur le volume L^3 dans la limite $L \to +\infty$. Formellement, on peut aussi décider de travailler par unité de volume, ce qui revient à remplacer partout L^3 par 1.

[24] En raison de l'isotropie, l'énergie du mode $(\vec{k}, \vec{\varepsilon})$ est indépendante de la direction de $\vec{k}$ et de la polarisation $\vec{\varepsilon}$, d'où la notation $\hbar\omega_k$.

les *opérateurs* de création et d'annihilation $a_{\vec{k}\vec{\varepsilon}}$ et $a^\dagger_{\vec{k}\vec{\varepsilon}}$, respectivement. Acceptant cette procédure, le Hamiltonien du champ libre se trouve muni de l'expression :

$$H_{\text{champ}} = \sum_{\vec{k},\varepsilon} \frac{1}{2}\hbar\omega_k \,(a_{\vec{k}\vec{\varepsilon}} a^\dagger_{\vec{k}\vec{\varepsilon}} + a^\dagger_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}}) \tag{25.57}$$

Sous cette forme, il est visible que les modes normaux du champ sont indépendants les uns des autres : H_{champ} est une somme sur l'énergie de *chaque* oscillateur (mode), forme caractéristique de l'absence de termes à deux ou plusieurs corps. On retrouve que le champ libre est effectivement une collection d'oscillateurs harmoniques sans interaction mutuelle[25].

Les opérateurs $a^\dagger_{\vec{k}\vec{\varepsilon}}$ et $a_{\vec{k}\vec{\varepsilon}}$ satisfont, mode à mode, les relations du type $[a, a^\dagger] = \mathbf{1}$ et commutent entre eux, d'où :

$$\boxed{[a_{\vec{k}\vec{\varepsilon}},\, a^\dagger_{\vec{k}'\vec{\varepsilon}'}] = \delta_{\vec{k}\vec{k}'}\,\delta_{\vec{\varepsilon}\vec{\varepsilon}'}\mathbf{1}} \tag{25.58}$$

ce qui permet de récrire l'expression (25.57) comme suit :

$$\boxed{H_{\text{champ}} = \sum_{\vec{k},\vec{\varepsilon}} \hbar\omega_k \, a^\dagger_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}} + \mathrm{C}^{\text{ste}}} \tag{25.59}$$

Formellement, la constante additive est ... aussi grande que l'on veut puisqu'elle est égale au nombre de modes distincts ; comme le produit $a^\dagger_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}} \stackrel{\text{déf}}{=} N_{\vec{k}\vec{\varepsilon}}$ est l'opérateur nombre d'excitations dans le mode $(\vec{k}, \vec{\varepsilon})$, l'état fondamental du champ (le *vide*) correspond à $N_{\vec{k}\vec{\varepsilon}} = 0$ quels que soient $\vec{k}$ et $\vec{\varepsilon}$, et on attribue conventionnellement à cet état une énergie *nulle*, étant entendu que seules importent les *différences* d'énergie. Une telle déclaration de principe est l'une des clés de voûte de la Théorie quantique des champs – c'est elle qui redonne au *champ* de Dirac toute son importance en le débarrassant d'emblée de toutes les difficultés liées à l'existence des états d'énergie négative (voir chapitre 20, sous-section 20.5.3). Formellement, ceci revient à omettre désormais la constante additive dans l'expression (25.59) de H_{champ}. Avec cette convention, tous les états excités du champ ont une énergie positive, le spectre étant par nature et par construction borné inférieurement par zéro.

L'identification physique des excitations du champ est maintenant claire : les deux opérateurs $a_{\vec{k}\vec{\varepsilon}}$ et $a^\dagger_{\vec{k}\vec{\varepsilon}}$ respectivement annihile et crée un *quantum* du champ électromagnétique d'impulsion $\hbar\vec{k}$, de polarisation $\vec{\varepsilon}$ et d'énergie $\hbar\omega_k \equiv \hbar kc$: ce quantum est évidemment un *photon*, dont la loi de dispersion $E_k = \hbar kc$ est typique d'une particule relativiste de masse nulle. L'état vide (fondamental du champ), $|\text{vac}\rangle$, est l'état contenant

[25]L'absence d'interaction entre les modes du champ permet de dire que les *photons* forment un *gaz parfait*. La matière (atomes, molécules, ...) est un intermédiaire obligé pour que deux photons puissent se voir ou avoir une filiation (ou être *corrélés*, voir EPR !) – d'où la nécessité d'un peu de matière pour assurer l'équilibre thermodynamique du champ, tout comme les collisions sont essentielles au maintien de l'équilibre d'un gaz parfait usuel (ou à sa relaxation vers l'état d'équilibre).

zéro photon, l'état $a^\dagger_{\vec{k}\vec{\varepsilon}}|\text{vac}\rangle$ est un état à un photon, etc. L'espace des états du champ est l'espace de Fock, tel qu'il a été défini dans la section 22.5.

L'expression (25.59) permet d'affirmer que les états propres du champ libre (en l'absence de matière) sont le produit tensoriel des états du type oscillateur harmonique, chaque oscillateur pouvant avoir une infinité de fréquences propres ω_k, selon la valeur de k :

$$|\text{état propre du champ}\rangle = |n_{\vec{k}_1\vec{\varepsilon}_1}\rangle \otimes |n_{\vec{k}_2\vec{\varepsilon}_2}\rangle \otimes \ldots = |n_{\vec{k}_1\vec{\varepsilon}_1} n_{\vec{k}_2\vec{\varepsilon}_2} \ldots\rangle \ , \qquad (25.60)$$

où tous les $n_{\vec{k}_i\,\vec{\varepsilon}_i}$ sont des entiers positifs ou nuls. Un état du champ est donc complètement déterminé par l'énoncé du nombre d'occupation $n_{\vec{k}_i\,\vec{\varepsilon}_i}$ du mode i ; les opérateurs de création et d'annihilation font varier d'une unité le nombre d'occupation relatif à leur mode :

$$a_{\vec{k}_i\vec{\varepsilon}_i}|\ldots n_{\vec{k}_i\vec{\varepsilon}_i}\ldots\rangle = \sqrt{n_{\vec{k}_i\vec{\varepsilon}_i}}\ |\ldots n_{\vec{k}_i\vec{\varepsilon}_i} - 1\ldots\rangle \ , \qquad (25.61)$$

et :

$$a^\dagger_{\vec{k}_i\vec{\varepsilon}_i}|\ldots n_{\vec{k}_i\vec{\varepsilon}_i}\ldots\rangle = \sqrt{n_{\vec{k}_i\vec{\varepsilon}_i} + 1}\ |\ldots n_{\vec{k}_i\vec{\varepsilon}_i} + 1\ldots\rangle \ , \qquad (25.62)$$

où tous les nombres cachés dans les ... sont inchangés d'un membre à l'autre. Enfin, conformément à (25.59), l'énergie propre correspondante est la simple somme :

$$H_\text{champ}|\ldots\, n_{\vec{k}_i\vec{\varepsilon}_i}\,\ldots\rangle = E_\text{champ}|\ldots\, n_{\vec{k}_i\vec{\varepsilon}_i}\ldots\rangle \ , \quad E_\text{champ} = \sum_i n_{\vec{k}_i\vec{\varepsilon}_i}\,\hbar\omega_{k_i} \ . \qquad (25.63)$$

Avec cette expression de H_champ, on trouve sans surprise que les opérateurs de création et d'annihilation satisfont les équations de Heisenberg :

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}a_{\vec{k}\vec{\varepsilon}\mathrm{H}} = \left[a_{\vec{k}\vec{\varepsilon}\mathrm{H}}, H_\text{champ}\right] = \omega_k a_{\vec{k}\vec{\varepsilon}\mathrm{H}} \ , \quad \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}a^\dagger_{\vec{k}\vec{\varepsilon}\mathrm{H}} = \left[a^\dagger_{\vec{k}\vec{\varepsilon}\mathrm{H}}, H_\text{champ}\right] = -\omega_k a^\dagger_{\vec{k}\vec{\varepsilon}\mathrm{H}} \qquad (25.64)$$

qui s'intègrent en :

$$a_{\vec{k}\vec{\varepsilon}\mathrm{H}}(t) = \mathrm{e}^{-\mathrm{i}\omega_k t}\, a_{\vec{k}\vec{\varepsilon}} \ , \qquad a^\dagger_{\vec{k}\vec{\varepsilon}\mathrm{H}}(t) = \mathrm{e}^{+\mathrm{i}\omega_k t}\, a^\dagger_{\vec{k}\vec{\varepsilon}} \ . \qquad (25.65)$$

Ces expressions sont bien sûr formellement identiques à celles obtenues pour les variables normales classiques introduites p. 1149.

À ce stade, il est possible de saisir l'importance de la quantification du champ et notamment le rôle crucial de ce que l'on appelle les *fluctuations du vide*, conséquences de la non-commutation entre les opérateurs du champ :

$$a_{\vec{k}\vec{\varepsilon}}a^\dagger_{\vec{k}\vec{\varepsilon}} \neq a^\dagger_{\vec{k}\vec{\varepsilon}}a_{\vec{k}\vec{\varepsilon}} \ . \qquad (25.66)$$

Soit un état du champ à nombre donné de photons d'impulsion et de polarisation bien définies, noté $|n_{\vec{k}\vec{\varepsilon}}\rangle$; la moyenne de la relation de commutation (25.58) dans cet état donne :

$$\langle n_{\vec{k}\vec{\varepsilon}}|a_{\vec{k}\vec{\varepsilon}}a^\dagger_{\vec{k}\vec{\varepsilon}}|n_{\vec{k}\vec{\varepsilon}}\rangle = 1 + \langle n_{\vec{k}\vec{\varepsilon}}|\, a^\dagger_{\vec{k}\vec{\varepsilon}}a_{\vec{k}\vec{\varepsilon}}|n_{\vec{k}\vec{\varepsilon}}\rangle = 1 + n_{\vec{k}\vec{\varepsilon}} \ . \qquad (25.67)$$

On peut deviner que le nombre de photons est directement relié à l'intensité des champs, laquelle contribue proportionnellement au nombre de transitions par unité de temps, autant en absorption qu'en émission induite[26]. C'est pourquoi le 1 au second membre de (25.67) fait toute la différence alors que l'autre terme, $n_{\vec{k}\vec{\varepsilon}}$, donne une parfaite symétrie entre absorption et émission induites. Le retour à l'équilibre thermique ne pouvant se faire que si le taux d'émission est *supérieur* au taux d'absorption[27], le 1 – signature de la non-commutation – est donc crucial pour le maintien de l'équilibre ou la relaxation vers l'équilibre thermique ; son apparition dans le formalisme constitue la justification de l'introduction phénoménologique par Einstein du coefficient A (voir chapitre 21, section 21.2). Bien sûr, dans la limite des très grands nombres d'occupation ($n_{\vec{k}\vec{\varepsilon}} \gg 1$), la non-commutation est *relativement* de peu d'importance.

Une fois la quantification effectuée, on peut écrire l'expression des opérateurs associés aux champs eux-mêmes. Par exemple, pour le champ électrique, il suffit de revenir à (25.40) ; effectuant la transformation de Fourier inverse pour obtenir l'expression dans l'espace réel, on trouve :

$$\vec{\mathcal{E}}(\vec{r}) = \mathrm{i} \sum_{\vec{\varepsilon}} \int \mathrm{d}^3k \sqrt{\frac{\hbar\omega}{16\pi^3\varepsilon_0}} \left(a_{\vec{k}\vec{\varepsilon}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} - a_{\vec{k}\vec{\varepsilon}}^\dagger \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\right) \vec{\varepsilon} \ , \tag{25.68}$$

dont le représentant de Heisenberg est :

$$\vec{\mathcal{E}}_{\mathrm{H}}(\vec{r}, t) = \mathrm{i} \sum_{\vec{\varepsilon}} \int \mathrm{d}^3k \sqrt{\frac{\hbar\omega_k}{16\pi^3\varepsilon_0}} \left(a_{\vec{k}\vec{\varepsilon}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \mathrm{e}^{-\mathrm{i}\omega t} - a_{\vec{k}\vec{\varepsilon}}^\dagger \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}} \mathrm{e}^{\mathrm{i}\omega t}\right) \vec{\varepsilon} \ . \tag{25.69}$$

25.1.3 Hamiltonien du système couplé charges + champ

En présence de matière (atomes, molécules,...) contenant des particules de masse m_i et de charges q_i, il convient de récrire les équations de Maxwell en y incluant la distribution de charge $\rho(\vec{r}) \stackrel{\text{déf}}{=} \sum_i q_i \delta(\vec{r} - \vec{r}_i)$ et le courant $\vec{j} \stackrel{\text{déf}}{=} \sum_i q_i \vec{v}_i \delta(\vec{r} - \vec{r}_i)$, $\vec{r}_i$ et $\vec{v}_i$ désignant la position et la vitesse de la particule d'indice i. La complication par rapport au champ libre est l'apparition d'une composante longitudinale $\vec{\mathcal{E}}_{\|}$ pour le champ électrique, et d'un potentiel scalaire U qui n'est plus identiquement nul ; toutefois, une fois adoptée la jauge de Coulomb, où le potentiel vecteur est transverse par définition, l'équation (25.44) montre que $\vec{\mathcal{E}}_{\|} = -\vec{\nabla} U$. Il est en outre possible de montrer [199] que U n'est autre que le potentiel de Coulomb :

$$U(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho(\vec{r}\,')}{\|\vec{r} - \vec{r}\,'\|} \mathrm{d}^3r' \ . \tag{25.70}$$

Il en résulte que le calcul de l'énergie totale du système couplé fait apparaître en plus le terme d'interaction électrostatique :

$$V_{\mathrm{C}} = \frac{1}{2} \frac{1}{4\pi\varepsilon_0} \sum_{i \neq j} \frac{q_i q_j}{\|\vec{r}_i - \vec{r}_j\|} \ ; \tag{25.71}$$

[26] Le qualificatif *induite* rappelle que c'est le champ qui provoque les transitions entre niveaux atomiques.

[27] Puisque la population thermique d'un état est une fonction décroissante de l'énergie.

pour des particules de vitesses faibles devant c, le Hamiltonien total s'écrit alors :

$$H = \sum_i \frac{1}{2} m_i \dot{\vec{r}}_i^2 + V_{\mathrm{C}} + H_{\mathrm{champ}} \ , \tag{25.72}$$

où H_{champ}, relatif au champ transverse libre, est le Hamiltonien (25.59). En effectuant maintenant le remplacement $m_i \vec{v}_i = \vec{p}_i - q_i \vec{A}(\vec{r}_i)$, on obtient :

$$\boxed{H = \sum_i \frac{1}{2} m_i [\vec{p}_i - q_i \vec{A}(\vec{r}_i)]^2 + V_{\mathrm{C}} + \sum_{\vec{k},\vec{\varepsilon}} \hbar\omega_k \, a_{\vec{k}\vec{\varepsilon}}^\dagger a_{\vec{k}\vec{\varepsilon}}} \tag{25.73}$$

qui est, en jauge de Coulomb, le Hamiltonien décrivant le système fermé matière et champ en interaction. Ce résultat achève le programme posé formellement dans l'écriture (25.1).

L'expression (25.73) deviendra opérationnelle dès que le potentiel vecteur $\vec{A}$ aura été exprimé à l'aide des opérateurs de champ. Par $\vec{\mathcal{E}} = -\partial_t \vec{A}$ et compte tenu de (25.69), on déduit le potentiel vecteur transverse :

$$\vec{A}_{\mathrm{H}}(\vec{r}, t) = \sum_{\vec{\varepsilon}} \int \mathrm{d}^3k \sqrt{\frac{\hbar}{16\pi^3 \varepsilon_0 \omega_k}} \left(a_{\vec{k}\vec{\varepsilon}}^\dagger \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}} \, \mathrm{e}^{\mathrm{i}\omega t} + a_{\vec{k}\vec{\varepsilon}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \, \mathrm{e}^{-\mathrm{i}\omega t} \right) \vec{\varepsilon} \ ; \tag{25.74}$$

de cette expression, on tire immédiatement le champ magnétique par $\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$. Noter que dans les égalités (25.68), (25.69) et (25.74), les opérateurs sont homogènes à $\mathrm{L}^{3/2}$ et satisfont $[a_{\vec{k}\vec{\varepsilon}}, a_{\vec{k}'\vec{\varepsilon}'}^\dagger] = \delta(\vec{k} - \vec{k}') \delta_{\vec{\varepsilon}\vec{\varepsilon}'}$. Lorsque l'on préfère utiliser les sommations discrètes dans l'espace réciproque, les expressions analogues sont :

$$\vec{\mathcal{E}}(\vec{r}) = \mathrm{i} \sum_{\vec{k},\vec{\varepsilon}} \sqrt{\frac{\hbar\omega_k}{2\varepsilon_0 L^3}} \left(a_{\vec{k}\vec{\varepsilon}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} - a_{\vec{k}\vec{\varepsilon}}^\dagger \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}} \right) \vec{\varepsilon} \ , \tag{25.75}$$

$$\vec{A}(\vec{r}) = \sum_{\vec{k},\vec{\varepsilon}} \sqrt{\frac{\hbar}{2\varepsilon_0 \omega_k L^3}} \left(a_{\vec{k}\vec{\varepsilon}}^\dagger \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}} + a_{\vec{k}\vec{\varepsilon}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \right) \vec{\varepsilon} \ . \tag{25.76}$$

Ici, les opérateurs sont sans dimension et satisfont $[a_{\vec{k}\vec{\varepsilon}}, a_{\vec{k}'\vec{\varepsilon}'}^\dagger] = \delta_{\vec{k}\vec{k}'} \delta_{\vec{\varepsilon}\vec{\varepsilon}'}$.

Ayant trouvé le potentiel vecteur quantique (25.74), l'interaction atome - champ réduite à sa plus simple et dominante expression[28] V_1 est :

$$V \simeq V_1 = -\frac{e}{\mu} \sum_{\vec{\varepsilon}} \int \mathrm{d}^3k \sqrt{\frac{\hbar}{16\pi^3 \varepsilon_0 \, \omega}} \left(a_{\vec{k}\vec{\varepsilon}}^\dagger \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}} + a_{\vec{k}\vec{\varepsilon}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \right) \vec{p}.\vec{\varepsilon} \ ; \tag{25.77}$$

alors, l'interaction V dans (25.1) est une combinaison *linéaire* des $a_{\vec{k}\vec{\varepsilon}}$ et des $a_{\vec{k}\vec{\varepsilon}}^\dagger$. Le processus élémentaire[29] entre états du champ à nombres donnés de photons est donc la

[28] V_1 est défini en (24.94).

[29] au sens où *une seule* interaction est en jeu.

variation d'une unité du nombre d'occupation – ce qui confirme que le photon est bien le *quantum* de champ –, cependant que l'atome lui-même change ou ne change pas[30] d'état (bien sûr, en allant à l'ordre n de la série de perturbation, on met en évidence des transitions possibles entre états du champ différant exactement de n photons, mais ces transitions procèdent par une succession de sauts élémentaires $\Delta n = \pm 1$). Ces mécanismes élémentaires peuvent être représentés par les célèbres diagrammes de Feynman.

Une fois l'expression des champs obtenue, examinons une autre conséquence de la non-commutation des opérateurs de champ, montrant bien en quoi cette non-commutation est responsable des fluctuations essentielles du champ, présentes même quand les champs sont nuls en moyenne. Dans l'état vide de photons, on a évidemment :

$$\langle \text{vac}|a^\dagger_{\vec{k}\vec{\varepsilon}}\, a_{\vec{k}\vec{\varepsilon}}|\text{vac}\rangle = 0 \ , \tag{25.78}$$

puisque $a_{\vec{k}\vec{\varepsilon}}|\text{vac}\rangle = 0$. L'équation (25.78) exprime le fait que dans l'état vide, le nombre *moyen* de photons est nul, quelles que soient l'impulsion et la polarisation (tous les modes sont vides en moyenne). Il en va d'ailleurs de même[31] de la *fluctuation* du nombre de photons, puisque $|\text{vac}\rangle$ est propre de l'opérateur $N_{\vec{k}\vec{\varepsilon}}$ et de son carré :

$$\langle \text{vac}|(a^\dagger_{\vec{k}\vec{\varepsilon}}\, a_{\vec{k}\,\vec{\varepsilon}})\,(a^\dagger_{\vec{k}\vec{\varepsilon}}\, a_{\vec{k}\,\vec{\varepsilon}})\,|\text{vac}\rangle = 0 \ . \tag{25.79}$$

Les champs $\vec{\mathcal{E}}$ et $\vec{\mathcal{B}}$ sont également nuls en moyenne dans tout état du type (25.60) et en particulier pour l'état vide (ils s'expriment en combinaison linéaire des $a^\dagger_{\vec{k}\vec{\varepsilon}}$ et des $a_{\vec{k}\,\vec{\varepsilon}}$ dont tous les éléments diagonaux sont nuls). En revanche, même dans l'état vide, les *carrés* de ces champs ne le sont pas ; en effet, l'élévation au carré d'une expression telle que (25.68) produit quatre types de termes :

$$a_{\vec{k}\vec{\varepsilon}}\, a_{\vec{k}\vec{\varepsilon}} \ , \qquad a^\dagger_{\vec{k}\vec{\varepsilon}}\, a^\dagger_{\vec{k}\vec{\varepsilon}} \ , \qquad a_{\vec{k}\vec{\varepsilon}}\, a^\dagger_{\vec{k}\vec{\varepsilon}} \ , \qquad a^\dagger_{\vec{k}\vec{\varepsilon}}\, a_{\vec{k}\vec{\varepsilon}} \ . \tag{25.80}$$

Tous ces termes, sauf ceux du troisième type, ont une valeur moyenne nulle dans l'état vide ; en vertu de la relation de commutation (25.58) :

$$\langle \text{vac}|a_{\vec{k}\vec{\varepsilon}}\, a^\dagger_{\vec{k}\,\vec{\varepsilon}}|\text{vac}\rangle = \langle \text{vac}|(\mathbf{1} + a^\dagger_{\vec{k}\,\vec{\varepsilon}}\, a_{\vec{k}\vec{\varepsilon}})|\text{vac}\rangle = 1 + 0 = 1 \ . \tag{25.81}$$

Ainsi, même dans l'état vide, les carrés des champs ne sont pas nuls ; ce sont ces fluctuations qui sont responsables, notamment, de l'émission spontanée, émission d'un photon alors que le champ est identiquement nul en moyenne – situation que l'on décrirait classiquement en disant *"le champ est nul"*. Ces fluctuations (localement infinies, voir problème 25.5.1) sont aussi responsables du *Lamb shift* (voir p. 898) – que l'on peut se représenter comme l'émission - absorption incessante de photons (virtuels) par un atome (voir ci-dessous).

---

[30] Si le photon est récupéré dans le milieu extérieur (processus réel), alors la conservation d'énergie exige que l'atome change d'état. Au contraire, un atome qui émet et réabsorbe un photon peut se retrouver dans son état de départ.

[31] La fluctuation du nombre de photons est d'ailleurs nulle dans tout état du type (25.60), pas seulement pour le vide.

♦ *Remarque*

Le Hamiltonien (25.73) ignore par construction tous les effets liés au spin. En pratique, il convient évidemment de rajouter *à la main* le couplage dipolaire magnétique $-\gamma_i \vec{S}_i.\vec{\mathcal{B}}$ entre un spin $\vec{S}_i$ et le champ magnétique. ♦

25.2 Description élémentaire de l'émission spontanée

À titre d'exemple, fixons-nous comme objectif de construire une description (très simplifiée) de l'émission spontanée. On admet de prime abord que l'atome possède seulement deux états, notés $|\mathrm{g}\rangle$ et $|\mathrm{e}\rangle$, en supposant bien sûr que l'élément de matrice de $\vec{p}$ entre ces deux états n'est pas nul[32]. En ce qui concerne le champ, on n'introduit que deux types d'états, l'état vide $|\mathrm{vac}\rangle$, et les états à un seul photon, $|\vec{k},\, \vec{\varepsilon}\rangle$. Au total, le système atome + champ possède 4 types d'états, obtenus par produit tensoriel :

$$\{|\mathrm{g},\, \vec{k},\, \vec{\varepsilon}\rangle\}_{\vec{k},\,\vec{\varepsilon}} \ , \quad |\mathrm{e},\, \mathrm{vac}\rangle \ , \quad |\mathrm{g},\, \mathrm{vac}\rangle \ , \quad \{|\mathrm{e},\, \vec{k},\, \vec{\varepsilon}\rangle\}_{\vec{k},\,\vec{\varepsilon}} \ ; \tag{25.82}$$

tous ces états sont propres de H_0 (atome et champ découplés), d'énergies respectives[33] :

$$\hbar kc \ , \qquad E_\mathrm{e} \ , \qquad 0 \ , \qquad E_\mathrm{e} + \hbar kc \ . \tag{25.83}$$

Dans le sous-espace engendré par les vecteurs (25.82), l'interaction atome - champ, V, est diagonale par blocs : il ne peut y avoir d'élément de matrice non-nul qu'entre deux états différant par un photon ; en outre, il faut en plus que $\vec{p}$ ait un élément de matrice également non-nul, ce qui exclut (par parité) tous les éléments de matrice diagonaux vis-à-vis des variables atomiques. Si on définit les projecteurs :

$$P \stackrel{\text{déf}}{=} |\mathrm{e},\, \mathrm{vac}\rangle\langle \mathrm{e},\, \mathrm{vac}| + \sum_{\vec{k},\,\vec{\varepsilon}} |\mathrm{g},\, \vec{k},\, \vec{\varepsilon}\rangle\langle \mathrm{g},\, \vec{k},\, \vec{\varepsilon}| \ , \qquad Q \stackrel{\text{déf}}{=} \mathbf{1} - P \ , \tag{25.84}$$

on voit que $PVQ = QVP = 0$. En outre, dans chacun des deux sous-espaces définis par P et Q, V a tous ses éléments diagonaux nuls.

L'émission spontanée est manifestement décrite par les processus mettant en jeu les états qui apparaissent dans le projecteur P, qui sont donc exclusivement considérés dans la suite. Parmi ces états, certains sont quasi-dégénérés, voire strictement dégénérés : ils peuvent donc être les points de départ et d'arrivée d'une transition réelle. L'autre sous-espace contient deux types d'états forcément séparés par une grande différence d'énergie au moins égale à E_e, qui ne peuvent de ce fait apparaître qu'en tant qu'intermédiaires

[32] Les éléments diagonaux de $\vec{p}$, pris sur les états atomiques, sont évidemment nuls puisqu'il s'agit d'états *liés*. Typiquement, $|\mathrm{g}\rangle$ désigne un état de type S (le fondamental par exemple) et $|\mathrm{e}\rangle$ un état de type P ; ce couple satisfait bien les règles de sélection E1 obtenues antérieurement (voir sous-section 24.5.1).

[33] L'énergie du fondamental de l'atome est prise comme origine.

dans des processus virtuels. En définitive, s'agissant de décrire le plus simplement possible l'émission spontanée – émission réelle d'un photon vers l'extérieur –, on ne retient que les états[34] :

$$\{|\mathrm{g},\, \vec{k},\, \vec{\varepsilon}\rangle\}_{\vec{k},\,\vec{\varepsilon}} \ , \qquad |\mathrm{e},\, \mathrm{vac}\rangle \ . \tag{25.85}$$

À cette approximation, le spectre de H_0 se compose de la demi-droite réelle s'étendant de 0 à $+\infty$ (le photon libre, avec k variant de 0 à $+\infty$) et du point isolé E_e. Reprenant le langage utilisé à propos de la résolvante[35], $G_0(z) \stackrel{\text{déf}}{=} (z\mathbf{1} - H_0)^{-1}$ a donc un pôle simple en E_e, noyé dans le *continuum* de photons libres.

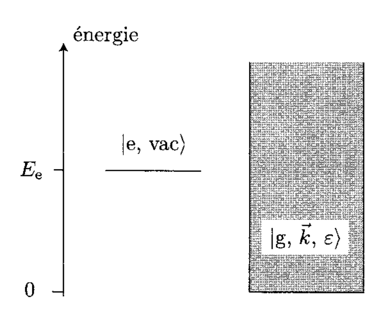


Figure 25.1: Schéma des niveaux d'énergie de H_0 dans l'espace effectif défini par P (voir (25.84)). Par comparaison avec la figure 22.3, il n'y a pas ici de *gap* entre l'état vide du champ et les états à un photon puisque celui-ci est de masse nulle.

La question que l'on se pose peut se formuler comme suit : le système étant préparé à $t = 0$ dans l'état $|\mathrm{e},\, \mathrm{vac}\rangle$ (pas de photon et atome à l'état excité), quelle est la probabilité pour l'atome d'être encore à l'état excité au temps $t > 0$? Cette probabilité dépend du temps puisque l'état préparé $|\mathrm{e},\, \mathrm{vac}\rangle$ n'est pas propre de H (mais seulement de H_0) : pour le Hamiltonien complet, c'est donc un état non-stationnaire et la probabilité cherchée est bien une fonction explicite du temps ; évidemment, la réponse à cette question donnera notamment la durée de vie du niveau excité. En d'autres termes, ce qu'il faut trouver c'est la variation au cours du temps de l'amplitude $A(t) \stackrel{\text{déf}}{=} \langle \mathrm{e},\, \mathrm{vac}|\Psi(t)\rangle$, dont le module au carré donnera la probabilité cherchée. En notant simplement $|E_\mathrm{e}\rangle \equiv |\mathrm{e},\, \mathrm{vac}\rangle$ et en reprenant les notations de la théorie des perturbations dépendant du temps (chapitre 24), l'expression de l'amplitude cherchée est :

$$A(t) = \langle E_\mathrm{e}|U(t)|\Psi(0)\rangle = \langle E_\mathrm{e}|U_0(t)U_1(t)|E_\mathrm{e}\rangle = \mathrm{e}^{\frac{1}{i\hbar}E_\mathrm{e}t}\langle E_\mathrm{e}|U_1(t)|E_\mathrm{e}\rangle \equiv \mathrm{e}^{\frac{1}{i\hbar}E_\mathrm{e}t}a(t) \ . \tag{25.86}$$

[34] L'espace effectif d'états ainsi défini a une dimension infinie puisque $\vec{k}$ prend ses valeurs dans $\mathbb{R}^3$, éventuellement discrétisé en gros points par les conditions cycliques habituelles.

[35] voir chapitre 23, sous-section 23.2.3.

Calcul perturbatif

La quantité essentielle est donc $a(t) = \langle E_\mathrm{e}|U_1(t)|E_\mathrm{e}\rangle$, dont il est facile d'exprimer la dérivée ; en effet :

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \langle E_\mathrm{e}|\frac{\partial}{\partial t}U_1(t)|E_\mathrm{e}\rangle \ , \tag{25.87}$$

et, compte tenu de :

$$\mathrm{i}\hbar\frac{\partial}{\partial t}U_1(t) = U_0^\dagger(t)\,V(t)\,U_0(t)\,U_1(t) \ \equiv\ V_\mathrm{I}(t)\,U_1(t) \ , \tag{25.88}$$

il vient :

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\langle E_\mathrm{e}|V_\mathrm{I}(t)U_1(t)|E_\mathrm{e}\rangle \ . \tag{25.89}$$

Faisons d'abord le calcul au plus bas ordre non-nul, soit en remplaçant U_1 par son expression au premier ordre[36] :

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\langle E_\mathrm{e}|V_\mathrm{I}(t)\Big[\mathbf{1} + \frac{1}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t'\,V_\mathrm{I}(t')\Big]|E_\mathrm{e}\rangle \ . \tag{25.90}$$

Dans le sous-espace asocié à P, V_I, tout comme V, est purement non-diagonal ; le terme $\mathbf{1}$ du crochet a donc une contribution nulle et il reste[37] :

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{1}{(\mathrm{i}\hbar)^2}\langle E_\mathrm{e}|V_\mathrm{I}(t)\int_0^t \mathrm{d}t'\,V_\mathrm{I}(t')|E_\mathrm{e}\rangle =$$
$$-\hbar^{-2}\int_0^t \mathrm{d}t'\,\langle E_\mathrm{e}|V_\mathrm{I}(t)\sum_{\vec{k},\,\vec{\varepsilon}}|g,\,\vec{k},\,\vec{\varepsilon}\rangle\langle g,\,\vec{k},\,\vec{\varepsilon}|V_\mathrm{I}(t')|E_\mathrm{e}\rangle \ . \tag{25.91}$$

En supposant le champ enfermé dans une très grande boîte, la somme discrète sur $\vec{k}$ porte sur des niveaux formant un *quasi-continuum* et peut être remplacée par une intégrale. En introduisant $E = \hbar kc$, la densité d'états $\rho(E)$ (qui incorpore la sommation sur les polarisations, soit $\sum_{\vec{k},\,\vec{\varepsilon}}\bullet \rightarrow \int \mathrm{d}E\rho(E)\bullet$) et en posant :

$$\langle \mathrm{g},\,\vec{k},\,\vec{\varepsilon}|V|E_\mathrm{e}\rangle = v(E) \qquad (E = \hbar kc) \ , \tag{25.92}$$

on voit que l'amplitude $a(t)$ satisfait :

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\frac{1}{2\pi\hbar}\int_0^{+\infty}\mathrm{d}E\int_0^t \mathrm{d}t'\,\Gamma(E)\,\mathrm{e}^{-(\mathrm{i}/\hbar)(E-E_\mathrm{e})(t-t')} \ , \tag{25.93}$$

où la fonction Γ, homogène à l'inverse d'un temps, est définie par :

$$\boxed{\Gamma(E) \stackrel{\text{déf}}{=} \frac{2\pi}{\hbar}\,\rho(E)\,|v(E)|^2} \tag{25.94}$$

[36] Au total, il s'agit d'un calcul au second ordre en V puisqu'une interaction est déjà présente en facteur de U_1 dans (25.89).

[37] Compte tenu de la forme de la perturbation, il faut bien *deux* interactions pour que l'élément de matrice diagonal sur $|E_\mathrm{e}\rangle$ soit différent de zéro.

Sans conteste, le second membre est réminiscent de la Règle d'or de Fermi (chapitre 24, éq.(24.87)), donnant la probabilité de transition par unité de temps. La fonction $\Gamma(E)$ est nulle pour $E < 0$ (le spectre de H_0 commence en $E = 0$), et tend vers zéro aux hautes énergies ; en effet, $v(E)$ contient en facteur l'élément de matrice $\langle \mathrm{e}|\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\vec{p}|\mathrm{g}\rangle \equiv \phi(\vec{k})$, qui est la transformée de Fourier d'un produit de deux fonctions représentant des états liés : leur extension spatiale linéaire est donc d'ordre a_0. Il en résulte que la fonction $\phi(\vec{k})$ est sensiblement non nulle seulement si $k \lesssim a_0^{-1}$, soit si $E \lesssim a_0^{-1}/(\hbar c)$, ou encore[38] $E \lesssim \alpha\mu c^2$. En d'autres termes, $\Gamma(E)$ a vaguement l'allure d'une cloche de largeur $\hbar\Omega \sim \alpha\mu c^2$. Par ailleurs, l'intégrale sur le temps t' fournit une partie réelle et une partie imaginaire, caractérisable chacune par une largeur d'ordre $\hbar/t$; si :

$$\frac{\hbar}{t} \ll \hbar\Omega \ , \tag{25.95}$$

alors on peut rejeter la borne t à $+\infty$ et on obtient[39] :

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\frac{1}{2\pi}\int_0^{+\infty} \mathrm{d}E\,\Gamma(E)\Big[\pi\,\delta(E_\mathrm{e} - E) + \mathrm{i}\,\mathcal{P}\frac{1}{E_\mathrm{e} - E}\Big] \tag{25.97}$$

où $\mathcal{P}$ désigne la partie principale de Cauchy. En posant maintenant[40] :

$$\boxed{\Delta(E) \stackrel{\text{déf}}{=} \frac{1}{2\pi}\mathcal{P}\int_0^{+\infty} \mathrm{d}E'\,\frac{\Gamma(E')}{E - E'}} \tag{25.98}$$

on obtient finalement :

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\,\Big[\frac{1}{2}\Gamma(E_\mathrm{e}) + \mathrm{i}\,\Delta(E_\mathrm{e})\Big] \tag{25.99}$$

qui s'intègre membre à membre en :

$$a(t) = 1 \,-\, \Big[\frac{1}{2}\Gamma(E_\mathrm{e}) + \mathrm{i}\,\Delta(E_\mathrm{e})\Big]\,t \ . \tag{25.100}$$

Il s'ensuit immédiatement que la probabilité $P(t) = |A(t)|^2 = |a(t)|^2 = a^*(t)a(t)$ est donnée au même ordre par :

$$P(t) = \big[1 \,-\, \frac{1}{2}\Gamma(E_\mathrm{e})t\big]^2 \,+\, \big[\Delta(E_\mathrm{e})\,t\big]^2 \simeq 1 - \Gamma(E_\mathrm{e})t + \mathcal{O}(t^2) \ . \tag{25.101}$$

Ce résultat, qui clôt le traitement perturbatif, n'est rien d'autre que la règle d'or de Fermi. En effet, la fonction $\Gamma(E_\mathrm{e})$ est égale à $|\mathrm{d}_t P(t)|$, et représente donc la probabilité de transition par unité de temps, ce que donne également la règle d'or, éq.(24.87), égalité dont le second membre coïncide bien avec la définition (25.94) de $\Gamma(E_\mathrm{e})$.

---

[38] α est la constante de structure fine.

[39] Se souvenir des relations :

$$\int_0^{+\infty} \mathrm{d}t\,\mathrm{e}^{-(\eta+\mathrm{i}x)t} = \frac{1}{\eta + \mathrm{i}x} = \frac{\eta}{x^2+\eta^2} - \mathrm{i}\,\frac{x}{x^2+\eta^2} \quad (\eta > 0) \ , \quad \lim_{\eta\to 0}\frac{1}{\eta + \mathrm{i}x} = \pi\,\delta(x) - \mathrm{i}\mathcal{P}\frac{1}{x} \ . \tag{25.96}$$

[40] La structure de l'expression donnant $\hbar\Delta(E)$ est celle d'une correction d'énergie au second ordre de la théorie des perturbations indépendantes du temps : somme de fractions ayant au numérateur le carré d'un élément de matrice et au dénominateur une différence d'énergie d'ordre zéro.

Clairement, (25.101) n'a de sens que si $\Gamma(E_e)t \ll 1$. Compte tenu de (25.95), il faut au total $\Omega^{-1} \ll t \ll \Gamma(E_e)^{-1}$, qui n'est possible que si :

$$\Omega^{-1} \ll \Gamma(E_e)^{-1} \quad \Longleftrightarrow \quad \Gamma(E_e) \ll \Omega \ ; \tag{25.102}$$

cette condition signifie que la largeur de l'état atomique *renormalisé*[41] doit être petite devant la largeur du *continuum* effectivement couplé à l'état atomique[42]. Ce que l'on vient d'obtenir est donc la description (approchée), valable aux temps courts, pour un niveau *étroit* couplé à un *continuum large.*

Par ailleurs, le calcul de l'amplitude $\langle \mathrm{g}, \vec{k}, \vec{\varepsilon}|U(t)|\mathrm{e}\rangle$ montre que celle-ci résonne à $k = k_e \stackrel{\text{déf}}{=} E_e/(\hbar c)$, ce qui permet de vérifier que l'énergie du photon final est égale à E_e, et exprime la conservation de l'énergie pour un processus réel.

Calcul non perturbatif

Il est possible de dépasser le traitement perturbatif et, notamment, de montrer que l'expression (25.101) est bien le début d'une exponentielle – alors $\Gamma(E_e)^{-1}$ pourra indiscutablement être identifié avec la durée de vie de l'état excité. En tout état de cause, le développement (25.101) n'a évidemment de sens que si $\Gamma(E_e)t \ll 1$, c'est-à-dire aux temps très courts, juste après l'excitation de l'atome, et c'est pourquoi il est vivement souhaitable de traiter le problème de façon plus complète.

La résolvante $G(z)$:

$$G(z) \stackrel{\text{déf}}{=} \frac{1}{z\mathbf{1} - H} \tag{25.103}$$

est un outil très commode pour dépasser le traitement perturbatif, comme on l'a déjà vu à propos de la théorie des perturbations indépendantes du temps[43]. On connaît le spectre de H_0, qui se compose de la demi-droite réelle $\mathbb{R}_+$ et du point isolé $E_e > 0$, noyé dans le *continuum* des photons libres ; en revanche, celui de H est inconnu à ce stade, mais il est bien évidemment situé dans la droite réelle. Considérons maintenant l'opérateur défini par l'intégrale :

$$U_+(t) \stackrel{\text{déf}}{=} \frac{1}{2\mathrm{i}\pi} \int_{\mathrm{C}_+} \frac{\mathrm{e}^{\frac{1}{\mathrm{i}\hbar} zt}}{z\mathbf{1} - H} \, \mathrm{d}z \tag{25.104}$$

où C_+ est une ligne parallèle à l'axe réel, située juste au-dessus et parcourue de $+\infty + \mathrm{i}0$ vers $-\infty + \mathrm{i}0$. Si $t < 0$, on referme le contour par un grand demi-cercle de rayon R du côté[44] $\Im z > 0$, qui n'apporte aucune contribution dans la limite $R \to +\infty$. Il n'y a aucune singularité dans le demi-plan $\Im z > 0$: par le théorème des résidus, l'intégrale est

[41] c'est-à-dire muni de sa largeur naturelle qui transforme la densité d'états atomique, la faisant passer d'une dent de Dirac à une *résonance* de largeur finie.

[42] Comme $\Gamma(E_e)$ contient l'interaction au carré, il est plausible que $\Gamma(E_e) \sim \alpha^2 \mu c^2 \sim \alpha\Omega \ll \Omega$, en général.

[43] Voir sous-section 23.2.3, et aussi le problème 21.6.5.

[44] $\Im z$ désigne la partie imaginaire de z.

nulle. Au contraire, pour $t > 0$, on referme par le bas et on ramasse toutes les singularités de la résolvante, qui constituent le spectre de H. Au total :

$$U_+(t) = \begin{cases} 0 & \text{si } t < 0 \\ e^{\frac{1}{i\hbar}Ht} & \text{si } t > 0 \end{cases} . \tag{25.105}$$

$U_+(t)$ est donc nul pour $t < 0$ et coïncide avec l'opérateur d'évolution[45] $U(t)$ pour $t > 0$. En terme d'intégrale sur l'axe réel, l'égalité suivante en résulte :

$$U_+(t) = \lim_{\varepsilon \to 0} \frac{1}{2i\pi} \int_{+\infty}^{-\infty} e^{\frac{1}{i\hbar}(E+i\varepsilon)t}\, G(E + i\varepsilon)\, dE \ , \tag{25.106}$$

où ε est un réel positif. Avec ces résultats, l'amplitude $A(t > 0) \stackrel{\text{déf}}{=} \langle E_e | U(t) | E_e \rangle$ possède la représentation intégrale :

$$A(t) = \frac{1}{2i\pi} \int_{C_+} G_e(z)\, e^{\frac{1}{i\hbar}zt}\, dz = \lim_{\varepsilon \to 0} \frac{1}{2i\pi} \int_{+\infty}^{-\infty} G_e(E + i\varepsilon)\, e^{\frac{1}{i\hbar}(E+i\varepsilon)t}\, dE \tag{25.107}$$

où $G_e(z) \stackrel{\text{déf}}{=} \langle E_e | G(z) | E_e \rangle$. Toute la question est alors de calculer $G_e(z)$, qui va se déduire de l'équation de Dyson $G(z) = G_0(z) + G_0(z)VG(z)$ (voir (23.138)) itérée une fois :

$$G(z) = G_0(z) + G_0(z)VG_0(z) + G_0(z)VG(z) \ . \tag{25.108}$$

G_0 est diagonal ; la perturbation V, au contraire, est purement non-diagonale et ne couple entre eux que des états différant d'un photon ; en prenant l'élément de matrice de (25.108) et en se référant à (25.92), il vient :

$$G_e(z) = \frac{1}{z - E_e} + 0 + \frac{1}{z - E_e} \int_0^{+\infty} \rho(E) v^*(E) \frac{1}{z - E} v(E)\, G_e(z)\, dE \ , \tag{25.109}$$

d'où l'on tire (voir (25.94)) :

$$G_e(z) = \frac{1}{z - E_e - \frac{\hbar}{2\pi} \int_0^{+\infty} \frac{\Gamma(E')}{z - E'}\, dE'} \ . \tag{25.110}$$

Compte tenu de la structure algébrique de cette expression – et du fait que $\Gamma(E) > 0$ (voir (25.94)) –, cette fonction ne peut avoir de pôle : par exemple, si $\Im z > 0$, l'intégrale du dénominateur a une partie imaginaire *négative* ; comme l'intégrale est précédée du signe $-$, la partie imaginaire du dénominateur est la somme de deux quantités positives et ne peut donc s'annuler ; la conclusion est la même si $\Im z < 0$. Le dénominateur ne peut donc jamais être nul[46] ; la fonction $G_e(z)$ est analytique dans tout le plan $\mathbb{C}$ privé du demi-axe réel $\mathbb{R}_+$.

---

[45] Cet opérateur, appelé propagateur avancé (ou causal), a été rencontré dans le Tome I, section 14.2, éq. (14.115).

[46] $G_e(z)$ n'a pas de pôle puisqu'un pôle est, par définition, une singularité isolée où le module de la fonction tend vers l'infini ; précisément, le dénominateur de $G_e(z)$ ne peut pas s'annuler.

Bien évidemment, $G_e(z)$ a heureusement d'autres singularités – voir le théorème de Liouville ! Ces singularités sont denses et définissent la coupure de 0 à $+\infty$ sur l'axe réel. En outre, le prolongement analytique de $G_e(z)$ dans le deuxième feuillet de Riemann possède un pôle au voisinage de l'aplomb de l'énergie atomique nue E_e, avec une partie imaginaire strictement négative – voir plus loin.

Les équations (25.110) et (25.98) donnent maintenant :

$$G_e(E + i0) = \frac{1}{E + i0 - E_e - \hbar\Delta(E) + i\frac{\hbar}{2}\Gamma(E)} \ ; \qquad (25.111)$$

de la même façon, on trouve :

$$G_e(E - i0) = \frac{1}{E - i0 - E_e - \hbar\Delta(E) - i\frac{\hbar}{2}\Gamma(E)} \ . \qquad (25.112)$$

Les deux expressions (25.111) et (25.112) confirment que la fonction $G_e(z)$ a une coupure s'étendant sur l'axe réel de 0 à $+\infty$ (la fonction $\Gamma(E)$ est identiquement nulle pour $E < 0$, de sorte que $G_e(E - i0) = G_e(E + i0)$ si $E < 0$) :

$$\boxed{G_{e\pm}(E) \stackrel{\text{déf}}{=} \lim_{\varepsilon\to 0} G_e(E \pm i\varepsilon) = \frac{1}{E - E_e - \hbar\Delta(E) \pm i\frac{\hbar}{2}\Gamma(E)}} \qquad (25.113)$$

Avant de continuer le calcul, il est utile d'interpréter ce résultat et d'élucider le contenu physique de la fonction $G_{e\pm}(E)$. Imaginons que le couplage avec le champ tende vers zéro. Alors :

$$G_{e\pm}(E) \ \rightarrow \ G^{(0)}_{e\pm}(E) = \frac{1}{E - E_e \pm i0} \ . \qquad (25.114)$$

En vertu de $\Im \lim_{\eta\to 0} \frac{1}{x + i\eta} = -\pi\delta(x)$, on en déduit :

$$\boxed{\mp \frac{1}{\pi} \Im G^{(0)}_{e\pm}(E) = \delta(E - E_e)} \qquad (25.115)$$

$\delta(E - E_e)$ n'est autre que la densité d'états de l'atome près de l'énergie E_e : quand le couplage atome - champ est nul, le seul état possible de l'atome (dit *nu*) est celui d'énergie exactement égale à E_e. Cette constatation permet de comprendre que la partie imaginaire de $G_{e\pm}(E)$ fournit la densité d'états de l'atome en présence du champ (atome *habillé*[47]) :

$$\rho_{at}(E) = \mp \frac{1}{\pi} \Im G_{e\pm}(E) = \frac{1}{2\pi} \frac{\hbar\Gamma(E)}{[E - E_e - \hbar\Delta(E)]^2 + \frac{1}{4}[\hbar\Gamma(E)]^2} \qquad (E \sim E_e) \ . \qquad (25.116)$$

En négligeant la variation (lente) de $\Delta(E)$ et de $\Gamma(E)$ au voisinage de E_e, la fonction $\rho_{at}(E)$ a essentiellement une variation lorentzienne, précurseur intuitif d'un déclin *exponentiel* de l'amplitude. Le lien entre fonction de Green et densité d'états a une validité universelle : d'une façon générale, la partie imaginaire de la trace de la *bonne* résolvante fournit la densité d'états complète.

Une fois l'élément de matrice $G_{e+}(E)$ déterminé, il suffit d'effectuer l'intégration (25.107) pour trouver l'amplitude $A(t)$, le long de C_+ ou de tout autre contour équivalent – un préalable à cette opération est l'identification des singularités de $G_e(z)$. En se

[47]expression due à Claude Cohen-Tannoudji ; en anglais, on dit *dressed atom*.

plaçant toujours dans l'hypothèse d'un couplage faible, ce calcul est grandement facilité en remplaçant la vraie fonction $G_\mathrm{e}(z)$ par une autre, dont les propriétés analytiques sont très différentes mais qui donnera à peu près la même valeur pour l'intégrale – au moins pour les temps pas trop longs. Pour ceci, on remarque que les fonctions Δ et Γ étant à variation lente, on ne commet pas une erreur grossière en remplaçant leur argument E par E_e. Ainsi apparaît la fonction[48] $G_{\mathrm{e}+}^{(\mathrm{ap})}(E)$:

$$G_{\mathrm{e}+}^{(\mathrm{app})}(E) \stackrel{\text{déf}}{=} \frac{1}{E - E_\mathrm{e} - \hbar\Delta_\mathrm{e} + \mathrm{i}\frac{\hbar}{2}\Gamma_\mathrm{e}} \ , \quad \Delta_\mathrm{e} \stackrel{\text{déf}}{=} \Delta(E_\mathrm{e}) \ , \quad \Gamma_\mathrm{e} \stackrel{\text{déf}}{=} \Gamma(E_\mathrm{e}) \ . \tag{25.117}$$

C'est ce que l'on appelle souvent *l'approximation du pôle* : une fois prolongée à vue, $G_{\mathrm{e}+}^{(\mathrm{app})}(z)$ a visiblement une et une seule singularité, un pôle en $E_\mathrm{e} + \hbar\Delta_\mathrm{e} - \mathrm{i}\frac{\hbar}{2}\Gamma_\mathrm{e}$, et a de ce fait, effectivement, des propriétés analytiques radicalement différentes de celles de la fonction réputée exacte $G_\mathrm{e}(z)$, laquelle a une coupure (et rien d'autre tant qu'elle n'est pas prolongée par continuité sur sa surface de Riemann). En tout cas, l'inversion de Fourier - Laplace (25.107) avec $G_{\mathrm{e}+}^{(\mathrm{app})}(z)$ est alors immédiate par résidus et donne :

$$\boxed{A(t) \simeq \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}[E_\mathrm{e}+\hbar\Delta_\mathrm{e}-\mathrm{i}\frac{\hbar}{2}\Gamma_\mathrm{e}]\,t} = \mathrm{e}^{\frac{1}{\mathrm{i}\hbar}(E_\mathrm{e}+\hbar\Delta_\mathrm{e})\,t}\,\mathrm{e}^{-\frac{1}{2}\Gamma_\mathrm{e}t} \qquad (t>0)} \tag{25.118}$$

Cette expression est transparente ; elle montre d'abord que, sous l'effet du couplage avec le champ, l'énergie de l'atome *nu* est déplacée de la quantité $\hbar\Delta_\mathrm{e}$, c'est ce que l'on appelle[49] le *Lamb shift*, que l'on peut se représenter comme le résultat de l'émission-absorption de photons virtuels par l'atome produisant une *renormalisation* de son énergie ; ces processus sont bien virtuels au sens où ils impliquent des états d'énergie arbitraire (voir la définition (25.98) de Δ_e, qui contient une somme sur l'énergie). Ce déplacement est loin d'être négligeable et joue un rôle important en Physique atomique[50] ; en particulier, il est responsable de la levée de dégénérescence des niveaux $2\mathrm{S}_{1/2}$ et $2\mathrm{P}_{1/2}$ de l'atome d'hydrogène, une fois *habillé* par les photons (voir p. 898). Chacun de ces niveaux a sa propre correction d'énergie mais il se trouve que cette correction est beaucoup plus grande[51] pour $2\mathrm{P}_{1/2}$ que pour $2\mathrm{S}_{1/2}$; au total, l'écart est manifeste et on trouve :

$$\boxed{E_{2\mathrm{S}_{1/2}} - E_{2\mathrm{P}_{1/2}} \simeq 1060\,\mathrm{MHz}} \tag{25.119}$$

Cette valeur est tout à fait comparable à l'élargissement Doppler à l'ambiante pour une raie optique et donne un effet très visible.

En outre, (25.118) donne :

$$\boxed{P(t) = |A(t)|^2 \simeq \mathrm{e}^{-\Gamma_\mathrm{e}t}} \tag{25.120}$$

[48] La densité d'états correspondante, $\rho_{\mathrm{at\ ap}}(E)$ est alors strictement lorentzienne.

[49] Cet effet a déjà été présenté dans le chapitre 20, p. 898, lors de la discussion de l'atome d'hydrogène traité en théorie de Dirac.

[50] Son calcul précis exige toutefois d'étendre le modèle présenté ici en détail.

[51] L'état $2\mathrm{S}_{1/2}$ est métastable – sa durée de vie est gigantesque, de l'ordre de la seconde – traduisant le fait que son couplage effectif au rayonnement est très faible (le retour vers le fondamental $1\mathrm{S}_{1/2}$ est une transition à deux photons, dont le principe fait l'objet du problème 24.7.4).

Dans l'approximation considérée, le déclin est effectivement purement exponentiel ; la durée de vie est $\Gamma_{\rm e}^{-1}$, la largeur naturelle est $\Gamma_{\rm e}$.

On voit que, dans l'approximation du pôle, tout se passe finalement comme si l'interaction atome - champ se bornait à remplacer l'énergie réelle $E_{\rm e}$ de l'atome sans champ par l'énergie *complexe* $\tilde{E}_{\rm e} = E_{\rm e} + \hbar\Delta_{\rm e} - {\rm i}\hbar\Gamma_{\rm e}/2$; le pôle réel de G_0 en $E_{\rm e}$ a donné naissance à un pôle situé à proximité mais ayant une partie imaginaire finie, et négative. Ceci est une vision simplifiée de ce qui se passe pour la fonction $G_{\rm e}(z)$; on sait qu'elle a une coupure, mais il est possible de la prolonger par continuité dans le deuxième feuillet de Riemann ; dans ce feuillet, le prolongement analytique a effectivement un pôle en $\tilde{E}_{\rm e}$. En tout cas, l'apparition, dans l'approximation du pôle, d'une énergie complexe fournit la justification théorique des approches phénoménologiques introduisant d'emblée par économie un Hamiltonien non-hermitique (voir l'exemple traité dans le problème 21.6.1).

Le traitement précédent repose sur un certain nombre d'approximations, notamment sur celle d'un couplage faible. Dans le cas contraire, des phénomènes intéressants se produisent, par exemple l'oscillation de Rabi, voir ci-après, section 25.3. Notons toutefois que même en restant à l'intérieur d'une hypothèse de couplage faible, le déclin n'est pas exponentiel jusqu'au bout : des arguments généraux[52] permettent de montrer que, pour des temps très grands devant $\Gamma_{\rm e}^{-1}$, la décroissance suit en fait une loi-puissance du genre $t^{-\lambda}$. Même si ce comportement non-exponentiel intervient très tard – au moment où il reste très peu d'atomes excités –, il n'en est pas pour autant sans intérêt : après tout, il constitue un exemple de curiosité[53] peut-être un peu inattendue, obtenue finalement dans un cadre physique assez banal.

25.3 De Wigner - Weisskopf à Rabi

La présentation précédente de l'émission spontanée décrit finalement la dynamique quantique d'un niveau discret (l'état $|E_{\rm e}\rangle$) à un *continuum* dans l'hypothèse du couplage faible[54]. La discussion qui suit ([203], [204]) décrit, toujours dans le cadre du modèle simple d'atome à deux niveaux, les modifications à apporter dans le cas où le couplage entre l'atome et le champ n'est soumis à aucune condition de ce type. On va voir que l'évolution temporelle de l'amplitude $A_{\rm e}(t)$ passe continûment d'un déclin exponentiel quand le *continuum* est large et le couplage faible à une oscillation harmonique sous-amortie réminiscente de l'oscillation de Rabi, quand le *continuum* est étroit et le couplage fort.

[52] À cause de la coupure, l'intégrale portant sur le prolongement analytique peut s'écrire comme une somme de deux termes : le résidu en $\tilde{E}_{\rm e}$ qui produit une composante de déclin exponentielle, et une intégrale (*lacet*) sur les deux bords de la coupure qui donne la décroissance algébrique en loi-puissance, $t^{-\lambda}$. Aux grands temps, l'exponentielle est éteinte et seule subsiste la décroissance en $t^{-\lambda}$, voir la note 6 p. 919 .

[53] Par exemple, si l'exposant λ est trop faible, on peut se trouver embarrassé pour définir une durée de vie au sens ordinaire (l'intégrale de $P(t)$ diverge si λ n'est pas supérieur à 1). En tout cas, il y a maintenant au moins deux échelles de temps : la durée de vie définie par $\Gamma_{\rm e}^{-1}$, c'est l'échelle courte, et une autre immensément longue par rapport à la précédente.

[54] Cette hypothèse intervient notamment pour justifier l'approximation du pôle.

On reprend l'atome à deux niveaux défini dans la section 25.2, couplé au champ avec les mêmes hypothèses. Pour faciliter la discussion, on introduit une constante de couplage sans dimension (c'est par exemple la constante de structure fine) en facteur de l'opérateur V. Ainsi, le Hamiltonien (25.1) s'écrit maintenant :

$$H = H_{\text{at}} + H_{\text{champ}} + gV \equiv H_0 + gV \ . \tag{25.121}$$

Avec cette modification, l'expression (25.110) – toujours valide – s'écrit :

$$G_{\text{e}}(z) = \frac{1}{z - E_{\text{e}} - g^2\,\Sigma(z)} \tag{25.122}$$

où la fonction :

$$\Sigma(z) \stackrel{\text{déf}}{=} \frac{\hbar}{2\pi}\int_0^{+\infty}\frac{\Gamma(E')}{z - E'}\,\mathrm{d}E' \ . \tag{25.123}$$

s'appelle *self-energy* et contient toutes les informations sur l'*habillage* du niveau atomique par les photons du champ[55].

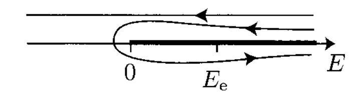


Figure 25.2: Déformation du contour d'intégration pour arriver à l'expression (25.124).

D'après (25.107), l'amplitude, à $t > 0$, est donnée par une intégrale le long d'un contour C_+ situé juste au-dessus de l'axe réel. En tirant vers le bas sur l'extrémité située en $-\infty + \mathrm{i}0$, on peut replier[56] le contour de façon à ce qu'il enlace la coupure, juste au-dessus et juste au-dessous de l'axe réel, en contournant le point de branchement en $E = 0$ (voir fig. 25.2). Dans ces conditions, l'amplitude $A_{\text{e}}(t)$ est donnée par :

$$A_{\text{e}}(t) = \frac{1}{2\mathrm{i}\pi}\int_{+\infty}^{0}[G_{\text{e}+}(E) - G_{\text{e}-}(E)]\mathrm{e}^{Et/(\mathrm{i}\hbar)}\,\mathrm{d}E \ , \tag{25.124}$$

où, conformément à (25.113) :

$$G_{\text{e}\pm}(E) = \frac{1}{E - E_{\text{e}} - g^2\Sigma_\pm(E)} \ , \qquad \Sigma_\pm(E) = \hbar[\Delta(E) \mp \mathrm{i}\,\Gamma(E)/2] \ . \tag{25.125}$$

[55]Pour un électron couplé aux phonons, on définit de façon analogue une *self-energy*, point de départ de la théorie du *polaron* – électron rhabillé dont la masse effective peut être si grande qu'il en devient quasi-immobile dans le cristal. Reprenant une image utilisée par Mattuck [181], l'habillage décrit par la *self-energy* évoque le nuage de poussière qui accompagne un cavalier au galop dans le désert.

[56]Ceci suppose que le couplage n'est pas trop fort au point de *détacher* un morceau du bas du *continuum*, produisant un pôle (point singulier isolé) d'énergie négative (voir p. 1172). On note que le petit cercle de contournement ne donne aucune contribution puisque l'intégrand est fini (mais non-analytique) en $E = 0$. De toute façon, on admettra quand il le faut que la fonction $\Gamma(E)$ tend vers zéro assez vite au voisinage du point de branchement ; remarquer d'ailleurs qu'une contribution non-nulle du point de branchement donnerait une probabilité $P(t)$ ne tendant pas vers zéro à l'infini.

Compte tenu de ceci, l'amplitude $A_e(t)$ est donnée par :

$$A_e(t) = \int_{-\infty}^{+\infty} \tilde{A}_e(E)\, e^{Et/(i\hbar)}\, dE \ , \tag{25.126}$$

où[57] :

$$\tilde{A}_e(E) = \frac{1}{\pi} \frac{g^2\hbar\Gamma(E)/2}{[E - E_e - g^2\hbar\Delta(E)]^2 + [g^2\hbar\Gamma(E)/2]^2} \ . \tag{25.127}$$

$\tilde{A}_e(E)$ ressemble à une lorentzienne, mais ses paramètres (centre et largeur) dépendent de la variable E. En réalité, la résonance se produit pour la (les) valeur(s) $\tilde{E}_e$ solution(s) de l'équation :

$$E - E_e - g^2\hbar\Delta(E) = 0 \ . \tag{25.128}$$

Au voisinage d'un tel zéro, $\tilde{E}_e$, on peut développer $\Delta(E)$ et $\Gamma(E)$ pour en déduire une expression simplifiée de $\tilde{A}_e$; un calcul élémentaire donne :

$$\tilde{A}_e(E) \simeq \frac{1}{1 - g^2\hbar\, \Delta'(\tilde{E}_e)} \frac{g^2\hbar\tilde{\Gamma}_e/(2\pi)}{(E - \tilde{E}_e)^2 + (g^2\hbar\tilde{\Gamma}_e/2)^2} \ , \tag{25.129}$$

où $\tilde{\Gamma}_e = \Gamma(\tilde{E}_e)$. $\tilde{E}_e$ est l'énergie renormalisée de l'état $|e\rangle$. Ainsi, au voisinage d'une résonance, $\tilde{A}_e$ est à peu près une lorentzienne ; le préfacteur dans (25.129), qui vient du calcul du résidu, donne le poids de cette résonance (l'assimilant à 1, on retrouve strictement l'approximation du pôle p. 1167).

Il n'est pas difficile de faire une discussion graphique permettant de préciser les zéros de l'équation (25.128). La fonction $\Gamma(E)$ est nulle pour $E < 0$, tend vers zéro en $E = 0$ et aux hautes énergies. Elle a donc un maximum quelque part, qui doit survenir dans les environs de E_e pour que le modèle soit significatif[58]. L'intégrale de $\Gamma(E)$ est une mesure globale de la force du couplage atome - champ ; il est donc naturel d'introduire la quantité Ω_1 définie comme :

$$\Omega_1^2 \stackrel{\text{déf}}{=} \frac{1}{2\pi\hbar} \int_0^{+\infty} \Gamma(E')\, dE' \ , \tag{25.130}$$

dont il est facile de voir qu'il s'agit d'une pulsation. Ces éléments étant précisés, on peut en dire un peu plus sur l'autre fonction importante, à savoir $\Delta(E)$ définie en (25.98). On note d'abord que quand $E \to \pm\infty$, on a :

$$\Delta(E) \simeq \frac{1}{2\pi} \frac{1}{E} \int_0^{+\infty} dE'\, \Gamma(E') = \frac{\hbar\Omega_1^2}{E} \ . \tag{25.131}$$

Δ s'annule quelque part, dans les environs de E_e en vertu de ce que l'on a supposé sur Γ ; en résumé, Δ a l'allure de la dérivée d'une résonance en cloche centrée vers E_e.

L'image graphique (voir fig. 25.3) de l'équation (25.128) est constituée par les points d'intersection de la courbe représentant la fonction $\hbar\Delta(E)$ avec la droite d'équation

[57] Dans (25.126), l'intégrale va en fait de 0 à $+\infty$, puisque la fonction $\Gamma(E)$ est nulle pour $E < 0$.

[58] Dans le cas contraire, ceci signifierait que le couplage au champ est intense pour des états éliminés d'emblée.

$g^{-2}(E - E_{\mathrm{e}})$, de pente d'autant plus raide que le couplage (mesuré par g) est petit. Pour le couplage faible, on a une seule intersection donnant une solution $\tilde{E}_{\mathrm{e}}$ très légèrement inférieure à l'énergie nue E_{e}. Quand le couplage augmente, cette racine décroît ; pour une valeur g_{c}, elle s'annule, cependant que, selon la forme précise de $\Delta(E)$, deux autres racines E_2 et E_3 sont déjà apparues ou vont apparaître. E_2 est située près de E_{e}, alors que E_3 est très nettement supérieure à E_{e}. Enfin, pour $g \gg g_{\mathrm{c}}$, E_2 est juste au-dessus de E_{e}, alors que E_3 est très grand devant l'écart $E_2 - E_{\mathrm{e}}$. La valeur précise de g_{c} s'obtient en écrivant que le premier membre de (25.128) est nul en $E = 0$, ce qui donne :

$$E_{\mathrm{e}} - g_{\mathrm{c}}^2\, \hbar\Delta(0) = 0 \quad \Longleftrightarrow \quad g_{\mathrm{c}} = \left[- \frac{E_{\mathrm{e}}}{\hbar\Delta(0)} \right]^{1/2} . \tag{25.132}$$

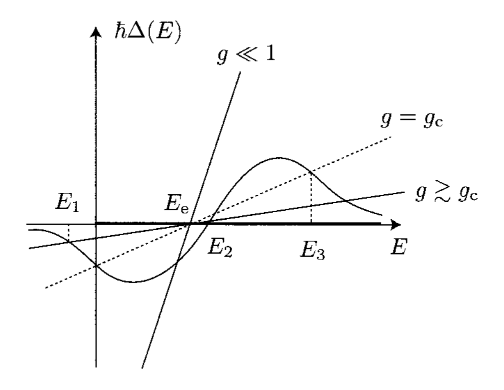


Figure 25.3: Discussion graphique de l'équation (25.128).

Discutons maintenant un peu en détail ces différents régimes. Pour le couplage faible (une seule racine près de E_{e}), il n'y a qu'une contribution à l'amplitude, approximativement donnée par (25.129), expression qui donne un régime purement exponentiel, avec une durée de vie $\tilde{\Gamma}_{\mathrm{e}}^{-1}$. En réalité, comme $\Gamma(E)$ s'annule pour $E < 0$, la dynamique révélera tôt ou tard le fait que l'amplitude est plus exactement donnée par (25.127). Cette différence implique la *queue* de la transformée de Fourier (près de $E = 0$) et modifiera donc le comportements aux temps longs. Un argument de *scaling* montre que si $\Gamma(E)$ s'annule comme E^{ν}, alors $A(t)$ varie comme $t^{-(\nu+1)}$ aux grands temps :

$$\Gamma(E) \sim E^{\nu} \ , E \rightarrow 0 \quad \Longleftrightarrow \quad A(t) \sim \frac{\mathrm{C}^{\mathrm{ste}}}{t^{\nu+1}} \qquad (t \rightarrow \infty) \ . \tag{25.133}$$

L'apparition d'une dépendance en loi-puissance traduit le fait que, $\tilde{A}_{\mathrm{e}}$ étant plus étroite qu'une pure Lorentzienne, sa transformée de Fourier est plus large en temps qu'une simple exponentielle.

Dans le cas du douplage fort ($g \gg g_{\mathrm{c}}$), il y a trois racines E_i ($i = 1, 2, 3$), dont on peut trouver les expressions approchées et les poids correspondants. E_1 est très négative : à force de *pousser* sur le bas du continuum, un morceau s'en est détaché ; comme

$\Gamma(E) = 0$ pour $E < 0$, cette résonance est de largeur nulle : il s'agit d'un état *discret*. Dans (25.128), on néglige E_e et on remplace $\Delta(E)$ par sa forme asymptotique (25.131), ce qui donne :

$$E_1 - g^2\hbar\frac{\hbar\Omega_1^2}{E_1} \simeq 0 \quad \Longleftrightarrow \quad E_1 \simeq -\hbar\Omega_1 \ . \tag{25.134}$$

Quant au poids, donné par $[1 - g^2\hbar\Delta'(E_1)]^{-1}$, on vérifie sans peine qu'il vaut à peu près 1/2. L'analyse procède de la même façon pour l'énergie E_3, que l'on trouve voisine de $+\hbar\Omega_1$, avec une largeur $\gamma_3 = (1/2)\Gamma(g\hbar\Omega_1)$ et un poids également voisin de 1/2. Enfin, on voit que $E_2 \simeq E_e$, le poids de cette résonance étant d'ordre $g^{-2} \ll 1$. En définitive, pour $g \gg 1$, $\tilde{A}_e$ se compose principalement, à poids égaux, d'un δ centré au voisinage de $-\hbar\Omega_1$ et d'une résonance très fine au voisinage de $+\hbar\Omega_1$. Dans le modèle en cours, il ne s'agit de rien d'autre que d'une oscillation de Rabi amortie, qui le sera d'autant plus lentement que la largeur de $\Gamma(E)$ est petite, rapportée à la bonne échelle (*continuum* étroit). Aux temps très longs devant γ_3^{-1}, il ne reste que la composante sur l'énergie voisine de $-\hbar\Omega_1$, d'où :

$$A(t) \simeq \frac{1}{2}\,\mathrm{e}^{+\mathrm{i}\,\Omega_1 t} \quad \Longleftrightarrow \quad |A(t)|^2 \simeq \frac{1}{4} \qquad (t \gg \gamma_3^{-1}) \ . \tag{25.135}$$

Au total, comme prévu, le couplage avec le *continuum* introduit une irréversibilité ("réduction du paquet d'ondes en énergie"), donnant une évolution non-unitaire quand on se concentre sur une partie des degrés de liberté (ici, la projection sur l'état atomique de départ). L'état final ne garde qu'un souvenir partiel de cet état initial.

Une dernière remarque, pour terminer. L'analyse précédente a certes été exprimée dans le langage de la Physique atomique, mais la méthode, les concepts et les outils utilisés ont une portée tout à fait générale et valent tout autant à chaque fois qu'il s'agit de décrire l'état instable d'un système (particule en Physique des hautes énergies – on parle souvent alors de *résonance* –, atome excité, quasi-particules en Matière condensée, complexe transitoire en Chimie, etc.).

♦

25.4 Description électrodynamique de l'interaction de van der Waals

L'interaction de van der Waals est bien connue : c'est l'interaction dominante entre deux systèmes *neutres* (atomes, molécules) situés à une distance R l'un de l'autre grande par rapport à la dimension $\sim a_0$ de chacun d'entre eux. Typiquement, elle prévaut sur des distances[59] de l'ordre de la dizaine d'Å et donne lieu à une *attraction* décrite par une énergie potentielle variant essentiellement comme $-R^{-6}$. Le but de cette section est de montrer que cette interaction est le résultat de l'échange de photons virtuels entre les atomes [200], et constitue un exemple élémentaire illustrant le fait général que toute

[59] C'est typiquement la distance moyenne entre atomes dans un gaz parfait situé dans les conditions normales. La liaison qui en résulte est très fragile : de l'ordre de quelques meV, bien inférieure à k_BT à l'ambiante.

interaction entre deux systèmes est médiée par un échange de particules véhiculant, en quelque sorte, l'information d'un système à l'autre. On trouvera tout naturellement que cette interaction n'est pas *instantanée* – elle est retardée –, quoique, dans les applications courantes, le retard soit en général tout à fait négligeable.

Soit deux atomes à un électron, dont les noyaux A et B, supposés infiniment massifs, se trouvent aux points $\vec{R}_A$ et $\vec{R}_B$; on note $\vec{R} = \vec{R}_B - \vec{R}_A$, $\vec{r}_\alpha$ et $\vec{r}_\beta$ les rayons vecteurs des électrons de l'un et de l'autre par rapport aux deux noyaux. En jauge de Coulomb, le Hamiltonien du système est :

$$H = \frac{1}{2m}[\vec{p}_\alpha - e\vec{A}(\vec{R}_A + \vec{r}_\alpha)]^2 + V_A + \frac{1}{2m}[\vec{p}_\beta - e\vec{A}(\vec{R}_B + \vec{r}_\beta)]^2 + V_B + V_{AB} + H_{\text{champ}} \ . \quad (25.136)$$

$\vec{A}$ est le potentiel vecteur transverse, H_{champ} est donné par (25.59). Les termes V_{AB} et V_I (I = A, B) désignent les interactions coulombiennes inter- et intra-atomiques respectivement ; en particulier :

$$V_{AB} = \frac{e'^2}{R} - \frac{e'^2}{\|\vec{R} - \vec{r}_\alpha\|} - \frac{e'^2}{\|\vec{R} + \vec{r}_\beta\|} + \frac{e'^2}{\|\vec{R} - \vec{r}_\alpha + \vec{r}_\beta\|} \ . \quad (25.137)$$

Par hypothèse, on a $\|\vec{r}_\alpha\| \sim \|\vec{r}_\beta\| \ll R$, ce qui autorise à faire un développement limité de cette expression ; un calcul sans difficulté conduit à :

$$V_{AB} = \frac{1}{4\pi\varepsilon_0 R^3}\Big(\vec{D}_\alpha.\vec{D}_\beta - 3\frac{\vec{D}_\alpha.\vec{R}}{R}\frac{\vec{D}_\beta.\vec{R}}{R} + \mathcal{O}\big(\frac{a_0}{R}\big)^3\Big) \ , \quad (25.138)$$

où $\vec{D}_\alpha = e\vec{r}_\alpha$ et $\vec{D}_\beta = e\vec{r}_\beta$ sont les moments dipolaires électriques des deux atomes. En théorie élémentaire, ce terme est traité par perturbation ; il n'y a pas de correction au premier ordre puisque, chaque état atomique ayant une parité déterminée, les éléments diagonaux des moments dipolaires sont nuls. La correction au second ordre, pour chaque atome dans l'état fondamental, est négative (voir chapitre 23, éq. (23.108)), et varie visiblement comme R^{-6} : c'est l'explication la plus simple de l'attraction de van der Waals.

L'idée qui va être développée est d'introduire une transformation S définissant une nouvelle représentation dans laquelle les atomes sont exclusivement couplés au rayonnement transverse. En faisant disparaître le terme statique (instantané) dipôle - dipôle (25.138), cette nouvelle représentation met en avant l'interaction entre atomes véhiculée par le champ électromagnétique, donc forcément *retardée*.

$\vec{A}(\vec{r})$ notant toujours le potentiel vecteur, soit la transformation unitaire $\mathcal{S}$:

$$\mathcal{S} \stackrel{\text{déf}}{=} e^{\frac{1}{i\hbar}[\vec{D}_\alpha.\vec{A}(\vec{R}_A) + \vec{D}_\beta.\vec{A}(\vec{R}_B)]} \ ; \quad (25.139)$$

l'expression (25.76) du potentiel vecteur permet d'établir que $[A_u(\vec{r}),\, A_{u'}(\vec{r}\,')] = 0$ quelles que soient les composantes u et u' de $\vec{A}$, et quels que soient $\vec{r}$ et $\vec{r}\,'$. En effet, la distribution des sommes $\sum_{\vec{k}\vec{\varepsilon}}\sum_{\vec{k}\,'\vec{\varepsilon}\,'}$ fait apparaître les termes :

$$\varepsilon_u\varepsilon'_{u'}e^{-i\vec{k}.\vec{r}}e^{i\vec{k}\,'.\vec{r}\,'}[a^\dagger_{\vec{k}\vec{\varepsilon}},\, a_{\vec{k}\,'\vec{\varepsilon}\,'}] + \varepsilon_u\varepsilon'_{u'}e^{i\vec{k}.\vec{r}}e^{-i\vec{k}\,'.\vec{r}\,'}[a_{\vec{k}\vec{\varepsilon}},\, a^\dagger_{\vec{k}\,'\vec{\varepsilon}\,'}] \ , \quad (25.140)$$

qui, en vertu de (25.58), sont égaux à $-\varepsilon_u\varepsilon'_{u'}\mathrm{e}^{-\mathrm{i}\vec{k}.(\vec{r}-\vec{r}\,')}+\varepsilon_u\varepsilon'_{u'}\mathrm{e}^{\mathrm{i}\vec{k}.(\vec{r}-\vec{r}\,')}$. Ils interviennent dans une somme sur le vecteur $\vec{k}\in\mathbb{R}^3$: quand $\vec{k}$ est changé en $-\vec{k}$, chaque vecteur $\vec{\varepsilon}$ est renversé ($\vec{\varepsilon}_1$, $\vec{\varepsilon}_2$ et $\vec{k}$ forment toujours un repère droit direct), ce qui ne fait rien au total. Il en résulte que, dans l'intégration sur $\vec{k}$, les deux termes se compensent d'une moitié de $\mathbb{R}^3$ à l'autre. La commutation de $\vec{A}$ avec lui-même entraîne que l'exponentielle de définition de la transformation $\mathcal{S}$ peut être factorisée pour donner $\mathcal{S}=\mathcal{S}_\mathrm{A}\mathcal{S}_\mathrm{B}$ avec $\mathcal{S}_\mathrm{I}\stackrel{\text{déf}}{=}\mathrm{e}^{\frac{1}{\mathrm{i}\hbar}\vec{D}_\alpha.\vec{A}(\vec{R}_\mathrm{I})}$, $\mathrm{I}=\mathrm{A},\,\mathrm{B}$.

Trouvons maintenant les transformés des variables dynamiques par la transformation $\mathcal{S}$. De toute évidence $\mathcal{S}_\mathrm{A}\vec{r}_\alpha\,\mathcal{S}_\mathrm{A}^\dagger=\vec{r}_\alpha$; quant à $\vec{p}_\alpha$, on a :

$$\mathcal{S}_\mathrm{A}\vec{p}_\alpha\,\mathcal{S}_\mathrm{A}^\dagger=\vec{p}_\alpha+\frac{e}{\mathrm{i}\hbar}\big[\vec{r}_\alpha.\vec{A}(\vec{R}_\mathrm{A}),\,\vec{p}_\alpha\big]+\Big(\frac{e}{\mathrm{i}\hbar}\Big)^2\big[\vec{r}_\alpha.\vec{A}(\vec{R}_\mathrm{A}),\,\big[\vec{r}_\alpha.\vec{A}(\vec{R}_\mathrm{A}),\,\vec{p}_\alpha\big]\big]+\ldots\;; \tag{25.141}$$

$\vec{p}_\alpha$ n'agit que sur $\vec{r}_\alpha$ ($\vec{R}_A$ est un simple paramètre et $\vec{A}(\vec{R}_A)$ ne contient que des opérateurs de champ) ; en raison de $[f(q),\,p]=\mathrm{i}\hbar f'(q)$, le premier terme du développement est donc $\frac{e}{\mathrm{i}\hbar}\,\mathrm{i}\hbar\vec{A}(\vec{R}_A)$; en conséquence, tous les commutateurs d'ordre supérieur sont nuls :

$$\mathcal{S}_\mathrm{A}\vec{p}_\alpha\,\mathcal{S}_\mathrm{A}^\dagger=\vec{p}_\alpha+e\vec{A}(\vec{R}_\mathrm{A})\equiv\vec{p}_\alpha-\mathrm{i}\hbar\,\mathcal{S}_\mathrm{A}\vec{\nabla}\mathcal{S}_\mathrm{A}^\dagger\;. \tag{25.142}$$

On en déduit que le transformé du premier terme cinétique de (25.136) est :

$$\frac{1}{2m}[\vec{p}_\alpha+e\vec{A}(\vec{R}_\mathrm{A})-e\vec{A}(\vec{R}_\mathrm{A}+\vec{r}_\alpha)]^2\simeq\frac{\vec{p}_\alpha^{\,2}}{2m}\;, \tag{25.143}$$

la dernière égalité venant de l'oubli du déphasage spatial du potentiel vecteur sur l'espace de l'atome A, ce qui suppose que seuls sont pertinents les vecteurs d'onde tels que $ka_0\ll 1$, une inégalité assurée largement dans le domaine optique[60]. Dans cette aproximation, le transformé de $H-H_\text{champ}$ est donc simplement :

$$\mathcal{S}(H-H_\text{champ})\mathcal{S}^\dagger=\frac{\vec{p}_\alpha^{\,2}}{2m}+\frac{\vec{p}_\beta^{\,2}}{2m}+V_\mathrm{A}+V_\mathrm{B}+V_\mathrm{AB}\;. \tag{25.144}$$

Il convient maintenant de trouver le transformé par $\mathcal{S}$ du Hamiltonien H_champ ; pour simplifier les notations, on pose :

$$\rho_{\vec{k}\vec{\varepsilon}\mathrm{A}}\stackrel{\text{déf}}{=}\frac{\mathrm{i}}{\hbar}\sqrt{\frac{\hbar}{2\varepsilon_0\omega_kL^3}}\,\vec{D}_\mathrm{A}.\vec{\varepsilon}\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}_\mathrm{A}}\;, \tag{25.145}$$

et une expression analogue $\rho_{\vec{k}\vec{\varepsilon}\mathrm{B}}$ pour l'atome B. Avec ces définitions, et compte tenu de (25.76), la transformation $\mathcal{S}_\mathrm{A}$ s'exprime comme l'exponentielle d'une somme, mais comme les différents modes commutent entre eux (voir (25.58)), on peut factoriser l'exponentielle de sorte que :

$$\mathcal{S}_\mathrm{A}=\prod_{\vec{k}\vec{\varepsilon}}\mathrm{e}^{\rho^\dagger_{\vec{k}\vec{\varepsilon}\mathrm{A}}a_{\vec{k}\vec{\varepsilon}}-\rho_{\vec{k}\vec{\varepsilon}\mathrm{A}}a_{\vec{k}\vec{\varepsilon}}a^\dagger_{\vec{k}\vec{\varepsilon}}}\;; \tag{25.146}$$

[60] C'est typiquement une approximation de type dipolaire électrique, déjà rencontrée dans la sous-section 24.5.1.

pour chaque mode :

$$\mathcal{S}_{\text{A}}\, a_{\vec{k}\vec{\varepsilon}}\mathcal{S}_{\text{A}}^{\dagger} = \text{e}^{\rho_{\vec{k}\vec{\varepsilon}\text{A}}^{\dagger} a_{\vec{k}\vec{\varepsilon}} - \rho_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}}^{\dagger}}\, a_{\vec{k}\vec{\varepsilon}}\, \text{e}^{-\rho_{\vec{k}\vec{\varepsilon}\text{A}}^{\dagger} a_{\vec{k}\vec{\varepsilon}} + \rho_{\vec{k}\vec{\varepsilon}\text{A}} a_{\vec{k}\vec{\varepsilon}}^{\dagger}} \ , \tag{25.147}$$

et de même pour $a_{\vec{k}\vec{\varepsilon}}^{\dagger} a_{\vec{k}\vec{\varepsilon}}$. Par la formule de Glauber, on a, au choix :

$$\text{e}^{\rho^{\dagger} a - \rho a^{\dagger}} = \left\{ \begin{array}{l} \text{e}^{-\frac{1}{2}|\rho|^2}\, \text{e}^{\rho^{\dagger} a} \text{e}^{-\rho a^{\dagger}} \\ \text{e}^{+\frac{1}{2}|\rho|^2}\, \text{e}^{-\rho a^{\dagger}} \text{e}^{\rho^{\dagger} a} \end{array} \right. \ , \tag{25.148}$$

d'où :

$$\text{e}^{\rho^{\dagger} a - \rho a^{\dagger}} a^{\dagger} \text{e}^{-\rho^{\dagger} a + \rho a^{\dagger}} = \text{e}^{-\frac{1}{2}|\rho|^2} \text{e}^{\rho^{\dagger} a} \text{e}^{-\rho a^{\dagger}} a^{\dagger} \text{e}^{+\frac{1}{2}|\rho|^2} \text{e}^{+\rho a^{\dagger}} \text{e}^{-\rho^{\dagger} a} = \text{e}^{\rho^{\dagger} a} a^{\dagger} \text{e}^{-\rho^{\dagger} a} \ ; \tag{25.149}$$

avec $[a,\, a^{\dagger}] = \mathbf{1}$, on a immédiatement $\text{e}^{\rho^{\dagger} a}\, a^{\dagger}\, \text{e}^{-\rho^{\dagger} a} = a^{\dagger} + \rho^{\dagger}$; prenant les hermitiques conjugués en vertu de $(\mathcal{S} a^{\dagger} \mathcal{S}^{\dagger})^{\dagger} = \mathcal{S} a \mathcal{S}^{\dagger}$, on en déduit $\text{e}^{\rho^{\dagger} a}\, a^{\dagger}\, \text{e}^{-\rho^{\dagger} a} = a + \rho$.

On peut maintenant expliciter $\mathcal{S} H_{\text{champ}} \mathcal{S}^{\dagger}$:

$$\mathcal{S} H_{\text{champ}} \mathcal{S}^{\dagger} = \sum_{\vec{k}\vec{\varepsilon}} \hbar\omega_k (a_{\vec{k}\vec{\varepsilon}}^{\dagger} + \rho_{\vec{k}\vec{\varepsilon}}^{\dagger})(a_{\vec{k}\vec{\varepsilon}} + \rho_{\vec{k}\vec{\varepsilon}}) \ , \tag{25.150}$$

où on a posé $\rho_{\vec{k}\vec{\varepsilon}} \overset{\text{déf}}{=} \rho_{\vec{k}\vec{\varepsilon}\text{A}} + \rho_{\vec{k}\vec{\varepsilon}\text{B}}$. Le développement des deux parenthèses produit des termes linéaires par rapport aux $\rho_{\vec{k}\vec{\varepsilon}}$, et des termes quadratiques. En ce qui concerne les termes linéaires, et revenant à la définition (25.145) des $\rho_{\vec{k}\vec{\varepsilon}\text{I}}$, on trouve sans peine qu'ils donnent $-\vec{D}_{\text{A}}\vec{\mathcal{E}}(\vec{R}_{\text{A}}) - \vec{D}_{\text{B}}\vec{\mathcal{E}}(\vec{R}_{\text{B}})$, où $\vec{\mathcal{E}}(\vec{r})$ est le champ électrique dont l'expression est donnée en (25.68).

Les termes quadratiques sont :

$$\sum_{\vec{k}\vec{\varepsilon}} \hbar\omega_k \big(\rho_{\vec{k}\vec{\varepsilon}\text{A}} \rho_{\vec{k}\vec{\varepsilon}\text{A}}^{\dagger} + \rho_{\vec{k}\vec{\varepsilon}\text{B}} \rho_{\vec{k}\vec{\varepsilon}\text{B}}^{\dagger} + \rho_{\vec{k}\vec{\varepsilon}\text{A}} \rho_{\vec{k}\vec{\varepsilon}\text{B}}^{\dagger} + \rho_{\vec{k}\vec{\varepsilon}\text{A}}^{\dagger} \rho_{\vec{k}\vec{\varepsilon}\text{B}}\big) \ ; \tag{25.151}$$

chaque somme impliquant les termes carrés est une certaine énergie, que l'on note $\epsilon(\text{A})$ et $\epsilon(\text{B})$. Le terme de droite vaut :

$$\sum_{\vec{k}\vec{\varepsilon}} \frac{1}{2\varepsilon_0 L^3} \left[(\vec{D}_{\text{A}}.\vec{\varepsilon})\,(\vec{D}_{\text{B}}.\vec{\varepsilon})\,\text{e}^{\text{i}\vec{k}.\vec{r}} + \text{h.c.}\right] \ ; \tag{25.152}$$

La somme sur $\vec{\varepsilon}$ porte sur deux vecteurs unitaires $\vec{\varepsilon}_{\lambda}$ perpendiculaires au vecteur d'onde $\vec{k}$. Les trois vecteurs $\vec{\varepsilon}_{\lambda}$ $(\lambda = 1,\, 2)$ et $\frac{\vec{k}}{k}$ formant un repère orthonormé, pour tout vecteur $\vec{D}$, on peut écrire $\vec{D} = (\vec{\varepsilon}_1.\vec{D})\vec{\varepsilon}_1 + (\vec{\varepsilon}_2.\vec{D})\vec{\varepsilon}_2 + (\vec{k}.\vec{D})\frac{\vec{k}}{k^2}$, de sorte que le produit scalaire $\vec{D}.\vec{D}\,'$ s'écrit aussi $\sum_{\vec{\varepsilon}_{\lambda}\perp\vec{k}} (\varepsilon_{\lambda}.\vec{D})(\varepsilon_{\lambda}.\vec{D}\,') + \frac{1}{k^2}(\vec{k}.\vec{D})(\vec{k}.\vec{D}\,')$. On en déduit, à l'envers :

$$\sum_{\vec{\varepsilon}_{\lambda}\perp\vec{k}} (\vec{\varepsilon}_{\lambda}.\vec{D})(\vec{\varepsilon}_{\lambda}.\vec{D}\,') = \vec{D}.\vec{D}\,' - \frac{1}{k^2}(\vec{k}.\vec{D})(\vec{k}.\vec{D}\,') \ . \tag{25.153}$$

C'est cette expression que l'on reporte dans (25.151) ; revenant aux grandeurs continues en $\vec{k}$, le terme (25.151) s'écrit (la somme sur $\vec{\varepsilon}$ est maintenant explicite) :

$$\frac{1}{2\varepsilon_0 L^3}\Big(\frac{L}{2\pi}\Big)^3\int \mathrm{d}^3k\Big\{\Big[\vec{D}_\mathrm{A}\vec{D}_\mathrm{B}-\frac{1}{k^2}(\vec{k}.\vec{D}_\mathrm{A}).(\vec{k}.\vec{D}_\mathrm{B})\Big]\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}+\mathrm{h.c.}\Big\}\ . \tag{25.154}$$

En introduisant les composantes cartésiennes $D_{\mathrm{I}u}$ des vecteurs, la quantité entre crochets est $\sum_{u,v=x,y,z} D_{\mathrm{A}u}D_{\mathrm{B}v}(\delta_{uv}-k^{-2}k_uk_v)$. L'intégration du terme en δ_{uv} donne[61] $(2\pi)^3\delta(\vec{R})$, une contribution que l'on peut ignorer puisque $\vec{R}$ n'est jamais le vecteur nul ; l'autre terme donne l'intégrale :

$$-\int_{\mathbb{R}^3}\frac{k_uk_v}{k^2}\,\mathrm{e}^{\mathrm{i}\vec{k}\vec{R}}\,\mathrm{d}^3k=\frac{\partial^2}{\partial R_u\partial R_v}\int_{\mathbb{R}^3}\frac{1}{k^2}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\,\mathrm{d}^3k\ . \tag{25.156}$$

Avec $\mathrm{d}^3k=k^2\mathrm{d}k\mathrm{d}\Omega$, l'intégrale s'explicite comme suit :

$$\int_{\mathbb{R}^3}\frac{1}{k^2}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\,\mathrm{d}^3k=\int_0^{+\infty}\mathrm{d}k\int_0^{\pi}\sin\theta\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,\mathrm{e}^{\mathrm{i}kR\cos\theta}=\int_0^{+\infty}\mathrm{d}k\,\frac{4\pi\sin kR}{kR}\ ; \tag{25.157}$$

l'intégrale en k résiduelle se calcule aisément par résidus et vaut $\frac{2\pi^2}{R}$. Il reste à calculer la dérivée partielle apparaissant en (25.156) ; on trouve sans peine :

$$\frac{\partial^2}{\partial R_u\partial R_v}\frac{1}{R}=-\delta_{uv}\frac{1}{R^3}+3\frac{R_uR_v}{R^5}\ ; \tag{25.158}$$

On arrive ainsi maintenant à l'expression finale[62] de (25.154) :

$$\frac{1}{\varepsilon_0}\frac{1}{4\pi}\sum_{u,v}D_{\mathrm{A}u}D_{\mathrm{B}v}\frac{1}{R^3}\Big(-\delta_{uv}+3\frac{R_uR_v}{R^2}\Big)\equiv-\frac{1}{4\pi\varepsilon_0}\Big[\vec{D}_\mathrm{A}.\vec{D}_\mathrm{B}-3\frac{(\vec{D}_\mathrm{A}.\vec{R})(\vec{D}_\mathrm{B}.\vec{R})}{R^2}\Big]\ . \tag{25.159}$$

La comparaison avec (25.138) montre que ceci est très exactement *l'opposé* du terme dominant de V_AB – lequel est visiblement invariant par la transformation $\mathcal{S}$.

En définitive, le Hamiltonien total transformé est :

$$\boxed{\mathcal{S}H\mathcal{S}^\dagger=\frac{\vec{p}_\alpha^{\,2}}{2m}+V_\mathrm{A}+\frac{\vec{p}_\beta^{\,2}}{2m}+V_\mathrm{B}-\vec{D}_\mathrm{A}\vec{\mathcal{E}}(\vec{R}_\mathrm{A})-\vec{D}_\mathrm{B}\vec{\mathcal{E}}(\vec{R}_\mathrm{B})+\sum_{\vec{k}\vec{\varepsilon}}\hbar\omega_k\,a_{\vec{k}\vec{\varepsilon}}^\dagger a_{\vec{k}\vec{\varepsilon}}} \tag{25.160}$$

La différence essentielle est que, dans cette nouvelle représentation, chaque atome interagit exclusivement avec le champ transverse, $\vec{\mathcal{E}}(\vec{R}_\mathrm{I})$, calculé au centre de l'atome. De toute évidence, l'interaction mutuelle des atomes est toujours présente, mais elle est

[61]On rappelle que :

$$\int_{\mathbb{R}^3}\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\,\mathrm{d}^3k=(2\pi)^3\,\delta(\vec{r})\ . \tag{25.155}$$

[62]Un facteur 2 apparaît car les deux quantités hermitiques conjuguées dans (25.154) donnent la même intégrale.

maintenant traduite par le fait que chaque atome interagit avec un *même* champ, lequel apparaît ainsi comme le *vecteur* obligé de son interaction avec l'autre[63]. On doit ainsi s'attendre à ce que, si à un certain instant l'un des atomes est dans un état excité et l'autre au fondamental, ce dernier ne peut être excité qu'au bout d'un temps de l'ordre de $\frac{R}{c}$, temps de vol d'un photon du premier atome au second. Ce fait majeur (au plan des principes) est démontré ci-dessous.

Avant d'établir ce résultat et pour compléter l'interprétation physique déjà partiellement construite, analysons la grandeur physique $\vec{E}(\vec{r}\,)$ devenue le champ transverse $\vec{\mathcal{E}}(\vec{r}\,)$ dans la nouvelle représentation ; par définition, on a ainsi :

$$\vec{E}(\vec{r}\,) \stackrel{\text{déf}}{=} \mathcal{S}^\dagger\, \vec{\mathcal{E}}(\vec{r}\,)\, \mathcal{S} \ . \tag{25.161}$$

Le calcul de $\vec{E}(\vec{r}\,)$ s'effectue exactement suivant les mêmes lignes que précédemment. Par sa définition, on a maintenant (avec toujours $\rho_{\vec{k}\vec{\varepsilon}} \stackrel{\text{déf}}{=} \rho_{\vec{k}\vec{\varepsilon}\mathrm{A}} + \rho_{\vec{k}\vec{\varepsilon}\mathrm{B}}$) :

$$\vec{E}(\vec{r}\,) = \mathrm{i}\sum_{\vec{k}\vec{\varepsilon}} \sqrt{\frac{\hbar\omega_k}{2\varepsilon_0 L^3}}\Big[(a_{\vec{k}\vec{\varepsilon}} - \rho_{\vec{k}\vec{\varepsilon}})\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} - (a^\dagger_{\vec{k}\vec{\varepsilon}} - \rho^\dagger_{\vec{k}\vec{\varepsilon}})\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\Big]\vec{\varepsilon} \ , \tag{25.162}$$

$\vec{E}(\vec{r}\,)$ est visiblement de la forme $\vec{\mathcal{E}}(\vec{r}\,) + \delta\vec{\mathcal{E}}_\mathrm{A}(\vec{r}\,) + \delta\vec{\mathcal{E}}_\mathrm{B}(\vec{r}\,)$. En procédant comme précédemment, on trouve sans difficulté que :

$$\delta\vec{\mathcal{E}}_\mathrm{A}(\vec{r}\,) = \frac{1}{\varepsilon_0 L^3}\sum_{\vec{k}}\Big(\vec{D}_\mathrm{A} - \frac{\vec{k}.\vec{D}_\mathrm{A}}{k^2}\,\vec{k}\Big)\,\mathrm{e}^{\mathrm{i}\vec{k}.(\vec{r}-\vec{R}_\mathrm{A})} \ , \tag{25.163}$$

et finalement, $\vec{r}$ étant *à l'extérieur*[64] de l'atome A, il vient :

$$\delta\mathcal{E}_{\mathrm{A}v}(\vec{r}\,) = -\frac{1}{4\pi\varepsilon_0}\frac{1}{\|\vec{r}-\vec{R}_\mathrm{A}\|^3}\sum_u D_{\mathrm{A}u}\Big[\delta_{uv} - 3\frac{(r_u - R_{\mathrm{A}u})(r_v - R_{\mathrm{A}v})}{\|\vec{r}-\vec{R}_\mathrm{A}\|^2}\Big] \ . \tag{25.164}$$

Ce résultat montre que $-\vec{D}_\mathrm{B}.\delta\vec{\mathcal{E}}_\mathrm{A}(\vec{R}_\mathrm{B})$ est précisément le terme dominant de V_AB.

Pour simplifier le formalisme, on suppose maintenant que chaque atome a un état fondamental $|\mathrm{g}\rangle$ et trois états excités $|\mathrm{e}_u\rangle$ dégénérés possédant une symétrie axiale le long de l'axe Ou, situés tous à l'énergie $\hbar\omega_\mathrm{a}$ au-dessus du fondamental ; par exemple, le fondamental est un état S, cependant que les états excités sont du type P, représentables, outre la fonction radiale R_{n1}, par Y_{10} et par les combinaisons linéaires $Y_{11} \pm Y_{1-1}$. Dans cet espace d'états effectif pour chaque atome, l'opérateur moment dipolaire – purement non-diagonal puisqu'il est impair – est tel que :

$$\vec{D}|\mathrm{g}\rangle = \sum_{u=x,y,z} d_u \vec{e}_u|\mathrm{e}_u\rangle \ , \qquad d_u \equiv e r_u \tag{25.165}$$

[63] Deux objets sans interaction directe mais chacun couplé à un même système sont indirectement couplés (penser à deux billes reliées chacune à une troisième par un ressort).

[64] Cette précision est utile, et provient du fait que l'on rencontre $\delta(\vec{r}-\vec{R}_\mathrm{A})$ dans le cours du calcul, tout comme on a rencontré $\delta(\vec{R}\,)$ dans le calcul p.1176.

où $\vec{e}_u$ est le vecteur unitaire de l'axe Ou.

Afin de mettre en évidence le temps de vol du photon d'un atome à l'autre, on suppose que, dans l'état initial, A est dans un état excité et B dans son fondamental, et on cherche en fonction du temps la probabilité de trouver la situation inverse : A au fondamental et B à l'état excité. Pour ce faire, soit plus précisément l'amplitude de probabilité[65] $A_{uv}(t)$:

$$A_{uv}(t) \stackrel{\text{déf}}{=} \mathrm{e}^{-\mathrm{i}\omega_{\mathrm{a}}t}\,\langle \mathrm{g_A e}_{v\mathrm{B}}|U_1(t)|\mathrm{e}_{u\mathrm{A}}\mathrm{g_B}\rangle \ , \tag{25.166}$$

où, selon (25.160), la perturbation V (indépendante du temps !) est :

$$V = -\vec{D}_{\mathrm{A}}\vec{\mathcal{E}}(\vec{R}_{\mathrm{A}}) - \vec{D}_{\mathrm{B}}\vec{\mathcal{E}}(\vec{R}_{\mathrm{B}}) \ . \tag{25.167}$$

Compte tenu de la forme du champ transverse (voir (25.68)), qui est *linéaire* dans les opérateurs de création et d'annihilation, il faut au moins *deux* interactions pour que l'amplitude de probabilité ne soit pas nulle. S'en tenant à l'ordre le plus bas ($n = 2$ dans la série de perturbation (24.27)), on peut ainsi écrire :

$$A_{uv}(t) \simeq A_{uv}^{(2)}(t) = (\mathrm{i}\hbar)^{-2}\mathrm{e}^{-\mathrm{i}\omega_{\mathrm{a}}t}\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\,\langle \mathrm{g_A e}_{v\mathrm{B}}|V_{\mathrm{I}}(t_1)V_{\mathrm{I}}(t_2)|\mathrm{e}_{u\mathrm{A}}\mathrm{g_B}\rangle \ . \tag{25.168}$$

À cet ordre, tout est calculable mais le calcul est un peu long (voir problème 25.5.4). On se borne ici à discuter le résultat auquel il conduit, à savoir :

$$\boxed{A_{uv}(t) \simeq A_{uv}^{(2)}(t) = \mathrm{i}\alpha c\,\mathrm{e}^{-\mathrm{i}\omega_{\mathrm{a}}t} r_u r_v g_{uv}(k_{\mathrm{a}},\,\vec{R})\,\big(t - \frac{R}{c}\big)\theta\big(t - \frac{R}{c}\big)} \tag{25.169}$$

où α est la constante de structure fine, $k_{\mathrm{a}} \stackrel{\text{déf}}{=} \omega_{\mathrm{a}}/c$, et où la fonction $g_{uv}(K,\,\vec{R})$ est[66] :

$$g_{uv}(K,\,\vec{R}) \stackrel{\text{déf}}{=} \mathrm{e}^{\mathrm{i}|K|R}\,h_{uv}(K,\,\vec{R}) \ . \tag{25.170}$$

avec :

$$h_{uv}(K,\,\vec{R}) \stackrel{\text{déf}}{=} |K|^3\Big[\frac{1}{|K|R}\big(\delta_{uv} - \frac{R_uR_v}{R^2}\big) - \frac{1}{(|K|R)^3}\big(\delta_{uv} - 3\frac{R_uR_v}{R^2}\big)\big(1 - \mathrm{i}|K|R\big)\Big] \ . \tag{25.171}$$

Ce résultat est valide pour $t \gg \frac{2\pi}{\omega_{\mathrm{a}}}$, une condition requise pour que les trois niveaux excités retenus par simplicité soient effectivement les seuls à jouer un rôle[67]. Supposant que la transition est dans le domaine optique, ceci veut dire que l'expression

[65]On reprend ici les notations du chapitre 24 où $U(t) = U_0(t)U_1(t)$.

[66]Pour $|K|R \ll 1$, $g_{uv}(K,\,\vec{R}) \simeq -R^{-3}(\delta_{uv} - 3\frac{R_uR_v}{R^2})$, qui caractérise le champ électrostatique instantané d'un dipôle. Au contraire, pour $|K|R \gg 1$, $g_{uv}(K,\,\vec{R}) \simeq |K|^2\frac{\mathrm{e}^{\mathrm{i}|K|R}}{R}(\delta_{uv} - \frac{R_uR_v}{R^2})$, qui correspond au champ rayonné à grande distance par un dipôle.

[67]Se souvenir de l'allure de la fonction $f(T,\omega)$ définie en (24.57), qui agit comme un *filtre*. Physiquement, une durée trop courte implique un spectre large, susceptible d'exciter d'autres niveaux que ceux explicitement considérés ici.

(25.169) suppose t nettement plus grand que 10^{-15} s, et qu'elle valide seulement les retards $\frac{R}{c} \gtrsim \frac{2\pi}{k_\text{a}c}$, soit $k_\text{a}R \gg 1$: le traitement ne vaut donc que pour des atomes *très* éloignés l'un de l'autre, mais c'est aussi seulement dans une telle situation que le retard est pertinent.

Le point sans doute le plus remarquable du résultat (25.169) est l'apparition de la fonction échelon-unité $\theta\big(t-\frac{R}{c}\big)$ qui montre clairement que l'atome initialement au fondamental ne peut passer dans l'état excité qu'au bout d'un temps *fini*, qui est précisément le temps de vol d'un photon d'un atome à l'autre ; une fois cet intervalle de temps écoulé, l'amplitude de probabilité croît linéairement, mais il faut se garder d'extrapoler en temps ce comportement, puisque le calcul a été limité au second ordre, l'ordre le plus bas non-nul. Bien sûr, si la mise en évidence de ce retard est fondamentale au plan des principes, celui-ci reste évidemment extrêmement faible dans les situations les plus courantes : pour deux atomes séparés[68] de 30 Å, $\frac{R}{c}$ vaut... 10^{-17} s.

♦ *Remarque*

Le traitement précédent peut être appliqué à un nombre quelconque d'atomes ou molécules. Des phénomènes intéressants surviennent quand ceux-ci sont disposés dans l'espace suivant un réseau régulier $\mathcal{B}$, supposé rigide[69] pour simplifier. En pareil cas, l'invariance du système par toutes les translations discrètes du réseau assure que tous les états propres de celui-ci pourront être paramétrés par un bon nombre quantique, $\vec{q}$, ayant toutes les caractéristiques d'un vecteur d'onde[70].

Supposons que chaque atome possède deux états, $|\text{g}\rangle$ et $|\text{e}\rangle$, satisfaisant les règles de sélection dipolaire électrique et donnant lieu à une transition située dans l'optique ($\lambda \sim 6\,000$ Å). L'état fondamental du réseau, $|0\rangle$, est celui où tous les atomes sont dans leur état fondamental. En décidant d'exciter un seul atome, on obtient N états distincts dégénérés ; en réalité, les atomes interagissent les uns avec les autres, et les vrais états propres du réseau sont de la forme :

$$|\vec{q}\rangle = \frac{1}{\sqrt{N}} \sum_{\vec{R}\in\mathcal{B}} \text{e}^{\text{i}\vec{q}.\vec{R}} |\vec{R}\rangle \ , \tag{25.172}$$

$|\vec{R}\rangle$ désignant l'état du réseau (d'ordre zéro) où tous les atomes sont dans leur état fondamental à l'exception de celui situé au site $\vec{R}$ qui est, lui, dans son état

[68] C'est la distance typique moyenne entre deux atomes de gaz dans les conditions normales.

[69] Dans la suite, le cas échéant, on supposera le cristal macroscopique mais *fini*, contenant $N \ggg 1$ atomes.

[70] Cette affirmation peut sembler bien expéditive. Elle sera expliquée au chapitre 30 à propos de la dynamique d'un électron dans un réseau, et constitue le fondement du théorème de Bloch. Ici, pour s'en convaincre, il suffit de réaliser que le Hamiltonien du réseau H_lat commute avec tous les opérateurs de translation discrète $\mathcal{T}(\vec{R})$ associés aux translations discrètes $T(\vec{R})$ superposant le réseau à lui-même ; ces opérateurs forment visiblement un groupe, lequel est abélien (on a vu au chapitre 17 que deux translations, finies ou non, commutent toujours). En conséquence, il existe des vecteurs propres communs à H_lat et à tous les $\mathcal{T}(\vec{R})$. En vertu de $\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}\,') = \mathcal{T}(\vec{R}+\vec{R}\,')$, les valeurs propres de $\mathcal{T}(\vec{R})$ sont de la forme $\text{e}^{\text{i}\vec{q}.\vec{R}}$, où $\vec{q}$ est un certain vecteur à composantes réelles. L'état (25.172) s'obtient en appliquant l'opérateur $\sum \text{e}^{\text{i}\vec{q}.\vec{R}}\mathcal{T}(\vec{R})$ à l'un quelconque des vecteurs de base de l'espace d'états effectif retenu (on a déjà utilisé ce procédé afin de construire automatiquement les états de bonne symétrie pour des particules identiques, voir chapitre 22, section 22.3).

excité $|e\rangle$. L'énergie de l'état $|\vec{q}\rangle$ est une certaine fonction $E(\vec{q})$, faiblement dispersée si l'interaction entre les atomes est petite devant la différence d'énergie d'un atome isolé, $\hbar\omega_a \stackrel{\text{déf}}{=} E_e - E_g$. Cette hypothèse est réalisée pour des solides de type *moléculaire* (gaz rares, molécules organiques,..., voir sous-section 29.2.2).

Dans ces conditions, le terme de couplage champ - matière, V, s'obtient en généralisation immédiate de celui qui figure dans (25.160), et s'écrit dans des notations évidentes :

$$V = -\sum_{\vec{R}\in\mathcal{B}} \vec{D}_{\vec{R}}.\vec{\mathcal{E}}(\vec{R}) \ . \tag{25.173}$$

Par hypothèse, $\vec{D}_{\vec{R}} = \vec{D}(|e_{\vec{R}}\rangle\langle g_{\vec{R}}| + |g_{\vec{R}}\rangle\langle e_{\vec{R}}|)$, où $\vec{D}$ est indépendant de $\vec{R}$, tous les sites du réseau étant équivalents et se déduisant les uns des autres par les translations discrètes de celui-ci. *A priori*, le champ couple directement l'état fondamental du réseau à tous les états $|\vec{q}\rangle$. Plus précisément, quand on remplace le champ électrique par son expression (25.68), l'élément de matrice $\langle 0, \vec{k}, \vec{\varepsilon}|V|\vec{q}\rangle$ fait intervenir des sommes du genre :

$$S \stackrel{\text{déf}}{=} \sum_{\vec{R}\in\mathcal{B}} e^{i(\vec{q}-\vec{k}).\vec{R}} \ . \tag{25.174}$$

Une telle somme reconstitue le produit de trois peignes de Dirac, et n'est non nulle que si $\vec{q}$ et $\vec{k}$ diffèrent d'un vecteur[71] $\vec{K}$ tel que $e^{i\vec{K}.\vec{R}} = 1$:

$$\vec{k} = \vec{q} + \vec{K} \ , \qquad e^{i\vec{K}.\vec{R}} = 1 \ ; \tag{25.175}$$

l'ordre de grandeur du plus petit vecteur $\vec{K}$ non nul de ce type est $\sim \frac{1}{a}$, où a est la dimension linéaire typique de la maille cristalline. Admettons que seuls les processus résonnants sont pertinents, c'est-à-dire que les photons jouant un rôle important sont ceux d'énergie $\sim \hbar\omega_a$, correspondant au vecteur d'onde optique $k_a \ll \frac{1}{a}$; dans ces conditions, seul le vecteur $\vec{K} = \vec{0}$ compte effectivement, et on trouve que l'interaction entre les photons d'impulsion $\hbar\vec{k}$ et les excitations élémentaires du réseau caractérisées par le vecteur $\vec{q}$ sont verrouillées par la règle de sélection :

$$\boxed{\vec{k} = \vec{q}} \tag{25.176}$$

qui s'interprète comme une loi conservation effective de l'impulsion : le réseau étant au fondamental, il peut absorber un photon d'impulsion $\hbar\vec{k}$ en créant une excitation élémentaire d'impulsion $\hbar\vec{q}$ avec $\vec{q} = \vec{k}$.

Ainsi, dans $\mathbb{R}^3$, chaque état de photon $|\vec{k}, \vec{\varepsilon}\rangle$ est couplé à *un* seul état du réseau, quand celui-ci est au fondamental et inversement : quand le réseau est excité en $|\vec{q}\rangle$, il ne *voit* qu'un seul photon (potentiel), celui de même impulsion. En définitive, $\vec{k}$ étant donné, on est en présence d'un système à deux niveaux, dont chaque état d'ordre zéro est caractérisé par un certain vecteur d'onde ; la diagonalisation de ce

[71] On verra plus tard (chapitre 29) que ces vecteurs $\vec{K}$ appartiennent eux-mêmes à un réseau, appelé réseau réciproque. L'argument conduisant à l'expression (25.174) sera repris et détaillé lors de l'étude de la diffraction (voir sous-section 29.3.2).

système à deux niveaux fournit des excitations hybrides stationnaires de la matière et du champ couplés, combinaisons linéaires de deux états $|0\rangle \otimes |\vec{k}, \vec{\varepsilon}\rangle$ et $|\vec{k}\rangle \otimes |\text{vac}\rangle$; on les appelle *polaritons*.

La situation change du tout au tout si le réseau est de dimensionnalité inférieure à 3. Pour fixer les idées, supposons qu'il s'agisse d'un réseau plan $\mathbb{Z}^2$. Dans ces conditions, la règle de sélection (25.176) est modifiée, puisque l'on peut ajouter à l'impulsion du photon n'importe quel vecteur $\vec{q}_\perp$ perpendiculaire au plan, tout en assurant la non-nullité de la somme S ; dans ces conditions, la loi de conservation de l'impulsion, (25.176), se transforme en :

$$\boxed{\vec{k} = \vec{q} + \vec{q}_\perp} \tag{25.177}$$

la toute première conséquence est maintenant qu'un état matériel de $\vec{q}$ donné est couplé à un *continuum* de photons. On a vu plus haut qu'un état discret (ici $|\vec{q}\rangle$) couplé à un *continuum* a nécessairement une durée de vie *finie* : la réduction de la dimensionnalité de la matière produit des polaritons *instables*. De façon imagée, on peut dire que quand matière et rayonnement sont enfermés dans la même boîte *tridimensionnelle*, ils sont obligés de cohabiter et s'arrangent pour fabriquer des états hybrides stables. Si au contraire la matière est confinée dans un plan (ou sur une ligne), le photon peut bel et bien s'échapper.

Pour recevoir un sens physique, la durée de vie finie de l'excitation plane doit bien sûr être comparée à la durée de vie typique de l'état excité d'un seul atome (isolé). Le calcul [205] montre que l'effet est important dans la mesure où le couplage ci-dessus réduit considérablement la durée de vie radiative de l'excitation $|\vec{q}\rangle$, au point de rendre dominant ce processus de désexcitation, par rapport à d'autres phénomènes de conversion d'énergie (*via* les vibrations du réseau par exemple, omises dans l'analyse ci-dessus mais inévitables dans un système réel). ♦

25.5 Exercices et problèmes

25.5.1 Quelques propriétés du champ libre

Le Hamiltonien du champ libre est l'opérateur H_{champ} donné en (25.59), dont on omet la constante additive au motif que seules importent les différences d'énergie. $N_p \stackrel{\text{déf}}{=} a_p^\dagger a_p$, $p \equiv (\vec{k}, \vec{\varepsilon})$, est l'opérateur nombre de particules dans le mode p, dont les valeurs propres sont notées n_p, $|n_p\rangle$ étant le *ket* propre correspondant.

1. Pourquoi les états de base du champ libre sont-ils de la forme :

$$|\psi\rangle \stackrel{\text{déf}}{=} |n_{p_1}\rangle \otimes |n_{p_2}\rangle \otimes ... \otimes |n_{p_r}\rangle \otimes ... \equiv |n_{p_1} n_{p_2} ...\rangle \text{ ?} \tag{25.178}$$

2. Exprimer un état tel que $|\psi\rangle$ à l'aide des opérateurs de création agissant sur l'état vide.

3. Soit $|\text{vac}\rangle$ l'état sans aucun photon (vide), $|0000...\rangle$.
   (a) Quelles sont les valeurs moyennes des champs $\vec{\mathcal{E}}(\vec{r})$ et $\vec{\mathcal{B}}(\vec{r})$ dans l'état vide ?
   (b) Quelles sont les valeurs moyennes des *carrés* des champs $\vec{\mathcal{E}}(\vec{r})$ et $\vec{\mathcal{B}}(\vec{r})$ dans l'état vide ?
   (c) Pour éclaircir la difficulté précédente, on définit un champ *à gros grain* obtenu en faisant la moyenne sur un petit domaine entourant un point donné de $\mathbb{R}^3$. Plus précisément, on pose par exemple :

$$\vec{\mathbb{E}}(\vec{r}) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} \varpi(r')\vec{\mathcal{E}}(\vec{r}+\vec{r}')\,\mathrm{d}^3r' \ , \tag{25.179}$$

   où $\varpi(r)$ est analogue à une densité de probabilité normalisée sensiblement non nulle pour $r \lesssim r_0$. Écrire le développement en ondes planes de $\vec{\mathbb{E}}(\vec{r})$ en y faisant figurer la transformée de Fourier $\phi(\vec{k}) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\varpi(r)\,\mathrm{d}^3r$; expliquer comment $\phi(\vec{k})$ agit comme une fonction de coupure pour les grands vecteurs d'onde $\gtrsim r_0^{-1}$, donc coupe également les hautes fréquences.

   Montrer que la fluctuation de $\vec{\mathbb{E}}(\vec{r})$ est maintenant finie, pourvu que $|\phi(\vec{k})|^2$ décroisse plus vite que k^4 à l'infini.
   (d) Soit $\vec{\mathbb{E}}(\vec{r}, t)$ le champ à gros grain dans l'image de Heisenberg. Écrire l'expression de sa fonction d'auto-corrélation dans l'état vide, $C_{EE}(t, t')$:

$$C_{EE}(t,t') \stackrel{\text{déf}}{=} \frac{1}{2}\langle\text{vac}|\vec{\mathbb{E}}(\vec{r},t).\vec{\mathbb{E}}(\vec{r},t') + \vec{\mathbb{E}}(\vec{r},t').\vec{\mathbb{E}}(\vec{r},t)|\text{vac}\rangle \ . \tag{25.180}$$

   Montrer que cette fonction est sensiblement différente de zéro à condition que $t - t'$ soit, en module, inférieur à $\tau_c \stackrel{\text{déf}}{=} r_0/c$, temps qui apparaît comme le temps de corrélation des fluctuations du vide[72].

4. En Mécanique quantique, l'état le plus général du champ est de la forme :

$$|\Psi(t)\rangle = \sum_{p_1p_2...} c_{p_1p_2...}(t)|p_1p_2...\rangle \ . \tag{25.181}$$

   (a) Comment s'expriment les $c_{p_1p_2...}(t)$ en fonction de leurs valeurs initiales ?
   (b) Soit l'état $|\Psi_1(t)\rangle \stackrel{\text{déf}}{=} (\cos\theta|\text{vac}\rangle + \sin\theta|1_{\vec{k},\,\varepsilon}\rangle)$. Quelles sont les valeurs moyennes du nombre de photons et du champ électrique dans l'état $|\Psi_1(t)\rangle$?
   (c) Expliquer en quoi les aspects ondulatoire et corpusculaire sont complémentaires.

[72] La question de savoir quelle est la bonne longueur physique r_0 est fort subtile. Il est clair que pour mesurer les fluctuations du vide, il faut utiliser une sonde ; si celle-ci est constituée de particules de masse m, la longueur qui s'impose est la longueur d'onde Compton. L'augmentation de la masse des particules met en évidence des fluctuations de plus en plus importantes : l'infini obtenu avec $r_0 = 0+$ doit bien être considéré comme une réalité physique incontournable. En un sens, tout comme une trajectoire brownienne (théorique) nulle part différentiable, le champ doit être considéré comme *rugueux* à toute échelle.

5. En Statistique quantique, et pour l'équilibre canonique, l'opérateur densité est $\rho(\beta) = Z^{-1}(\beta)\mathrm{e}^{-\beta H_{\text{champ}}}$.

   (a) Expliquer pourquoi $\rho(\beta)$ a la forme factorisée suivante :

$$\rho(\beta) = \rho_{p_1}(\beta) \otimes \rho_{p_2}(\beta) \otimes ...\rho_{p_r}(\beta) \otimes ... \quad . \tag{25.182}$$

   (b) Calculer la valeur moyenne de N_p dans l'état d'équilibre canonique, $\langle N_p \rangle$, ainsi que sa fluctuation mesurée par l'écart quatratique ΔN_p^2. Montrer que ΔN_p^2 s'exprime très simplement en fonction de $\langle N_p \rangle$.

   (c) Connaissant $\langle N_p \rangle$, comment retrouve-t-on immédiatement la loi de Planck ?

   (d) Quelles sont les valeurs moyennes des champs dans l'état d'équilibre canonique ?

25.5.2 Hamiltonien spin - boson

Afin de décrire schématiquement l'interaction[73] entre un atome à deux niveaux et un certain mode du champ électromagnétique, on pose le Hamiltonien suivant :

$$H = \frac{\hbar\omega_0}{2}\sigma_z + \hbar\omega a^\dagger a + \hbar\Omega(a^\dagger\sigma_- + a\sigma_+) \equiv H_{\text{at}} + H_{\text{f}} + V \quad . \tag{25.183}$$

ω_0 et ω sont positifs, les σ_z, $\sigma_\pm$ sont les matrices de Pauli agissant dans l'espace des deux états atomiques notés $|\pm\rangle$:

$$H_{\text{at}}|\pm\rangle = \pm\frac{\hbar\omega_0}{2}|\pm\rangle \quad . \tag{25.184}$$

Les a et $a^\dagger$ sont les opérateurs de d'annihilation et de création usuels, relatifs au mode du champ considéré. Les états du champ libre (un seul mode) sont notés $|n\rangle$ ($n \in \mathbb{N}$) :

$$H_{\text{f}}|n\rangle = n\hbar\omega|n\rangle \quad . \tag{25.185}$$

Les états du système (atome + champ) sans interaction sont les produits tensoriels notés $|\pm, n\rangle$ et constituent une base de l'espace des états $\mathcal{E}$ pour le système couplé. L'action des opérateurs sur ces états est définie comme d'habitude ; par exemple :

$$\sigma_x|+, n\rangle = |-, n\rangle \ , \ a|\pm, n\rangle = \sqrt{n}|\pm, n-1\rangle \ , \ \sigma_z a^\dagger|-, n\rangle = -\sqrt{n+1}|-, n+1\rangle \ , \quad \text{etc} \ . \tag{25.186}$$

Le but du problème est de trouver la transformation unitaire U permettant d'engendrer les états propres de H à partir des $\{|\pm, n\rangle\}$.

1. En supposant $\omega < \omega_0$ pour fixer les idées, faire un schéma énergétique représentant l'état fondamental et les premiers états excités du système sans interaction.

[73] L'atome à deux niveaux est isomorphe à un spin $S = \frac{1}{2}$; le champ électromagnétique, représenté par les opérateurs a et $a^\dagger$ satisfaisant $[a, a^\dagger] = \mathbf{1}$, est un champ de bosons (photons), d'où la terminologie *spin - boson*.

2. Calculer $V|+,\, n\rangle$ et $V|-,\, n+1\rangle$ et illustrer les résultats par deux diagrammes décrivant l'action de V (on représentera l'atome par un trait continu et un photon en plus ou en moins par un trait ondulé).

3. Sans faire de calculs, expliquer pourquoi le Hamiltonien H est en fait une somme (directe) d'opérateurs H_n dont chacun agit exclusivement dans un sous-espace $\mathcal{E}_n$ de dimension égale à 2. Quel est l'état fondamental de H ?

4. Retrouver le résultat de la question précédente en argumentant à partir de la valeur du commutateur $[H, \mathcal{N}]$ où $\mathcal{N} \stackrel{\text{déf}}{=} \frac{1}{2}\sigma_z + a^\dagger a$.

5. Une base de $\mathcal{E}_n$ est constituée par les deux vecteurs $|+,\, n\rangle$ et $|-,\, n+1\rangle$. Écrire la matrice de H_n sur cette base ainsi ordonnée. Dans la suite, Σ_u désigne une matrice identique à la matrice de Pauli σ_u mais agissant cette fois dans l'espace $\mathcal{E}_n$. Décomposer la matrice de H_n sur les σ_u en déterminant ε_{0n} et $\vec{\varepsilon}_n$ dans l'écriture ci-après :
$$H_n = \varepsilon_{0n}\mathbf{1}_2 + \vec{\varepsilon}_n.\vec{\Sigma} \ . \tag{25.187}$$

6. Soit une transformation (unitaire) U_n telle que :
$$H_n = \varepsilon_{0n}\mathbf{1}_2 + \frac{\hbar\tilde{\omega}_n}{2}\, U_n\Sigma_z\, U_n^\dagger \tag{25.188}$$
où la pulsation $\tilde{\omega}_n$ (réelle positive) est pour l'instant indéterminée. Montrer que les deux vecteurs :
$$|\psi_{n\,+}\rangle \stackrel{\text{déf}}{=} U_n|+,\, n\rangle \ , \qquad |\psi_{n\,-}\rangle \stackrel{\text{déf}}{=} U_n|-,\, n+1\rangle \tag{25.189}$$
sont propres de H_n. Quelles sont les valeurs propres correspondantes ?

7. Il s'agit maintenant de trouver effectivement U_n. Pour ceci, on pose :
$$U_n = \exp\Big[\frac{1}{2}\theta_n(|+,\, n\rangle\langle -,\, n+1| + \lambda|-,\, n+1\rangle\langle +,\, n|)\Big] \equiv \mathrm{e}^{\frac{1}{2}\theta_n\xi_n} \ , \tag{25.190}$$
où θ_n est réel ; comment peut-on *a priori* choisir le scalaire λ ? Ce choix étant fait, exprimer U_n en fonction de Σ_y et l'écrire sous la forme d'une combinaison linéaire de Σ_y et de la matrice identité.

8. En déduire l'angle θ_n (par sa tangente) et l'expression de $\tilde{\omega}_n$.

9. Exprimer à l'aide de θ_n les vecteurs propres de H_n, $|\psi_{n\,+}\rangle$ et $|\psi_{n\,-}\rangle$, en combinaison linéaire des $|+,\, n\rangle$, $|-, n+1\rangle$ $(n \geq 0)$.

10. La transformation globale U est le produit des transformations U_n ; expliquer pourquoi on peut aussi l'écrire sous la forme $U = \mathrm{e}^{\frac{1}{2}\sum_n \theta_n\xi_n}$.

25.5.3 États cohérents du champ

On passe ici en revue quelques propriétés [200] des états du champ construits avec les états cohérents définis à propos de l'oscillateur harmonique (voir Tome I, chapitre 16, section 16.5).

1. Soit l'opérateur $S(\alpha) \stackrel{\text{déf}}{=} \mathrm{e}^{\alpha a^\dagger - \alpha^* a}$ où $\alpha \in \mathbb{C}$, où a et $a^\dagger$ sont relatifs à un certain mode du champ. L'opérateur $S(\alpha)$ est-il unitaire ?

2. À l'aide de la formule de Glauber (Tome I, éq. (13.215)), calculer le résultat de l'action de $S(\alpha)$ sur le vide, $|0\rangle$.

3. On note désormais $|\alpha\rangle \stackrel{\text{déf}}{=} S(\alpha)|0\rangle$.
   (a) Quelle est la probabilité P_n de trouver n photons dans l'état $|\alpha\rangle$? Comment s'appelle cette distribution de probabilité ?
   (b) Calculer le nombre moyen de photons dans l'état $|\alpha\rangle$, $\langle N\rangle$, et l'écart quadratique ΔN^2.
   (c) Combien valent $\langle N\rangle$ et ΔN^2 pour la distribution de Bose ?
   (d) Comparer la distribution $\{P_n\}_n$ et la distribution de Bose.
   (e) Le cas $N \gg 1$ correspond à la limite presque classique (très grands nombres quantiques). Comment se comparent dans ce cas les fluctuations relatives des deux distributions ?

4. Soit $|\Psi\rangle$ un certain état développé sur les états cohérents, $|\Psi\rangle = \frac{1}{\pi}\int \mathrm{d}^2\alpha|\alpha\rangle\langle\alpha|\Psi\rangle$, avec lequel on forme l'opérateur densité $\rho \stackrel{\text{déf}}{=} |\Psi\rangle\langle\Psi|$. Sous réserve d'existence, ρ admet la décomposition :

$$\rho = \int P(\alpha)|\alpha\rangle\langle\alpha|\mathrm{d}^2\alpha \ , \tag{25.191}$$

   où $\mathrm{d}^2\alpha \equiv \mathrm{d}\alpha_1\mathrm{d}\alpha_2$ quand $\alpha = \alpha_1 + \mathrm{i}\alpha_2$.
   (a) Montrer que la fonction $P(\alpha)$ est à valeurs réelles. Soit l'opérateur[74] $(a^\dagger)^n a^m$; former l'expression de sa moyenne dans l'état pur ρ.
   (b) Peut-on interpréter $P(\alpha)$ comme la probabilité d'observer l'état $|\alpha\rangle$?
   (c) Se référant à un oscillateur matériel, a est une combinaison linéaire de x et p ; expliquer en quoi ceci explique l'impossibilité précédente.

[74] Un produit d'opérateurs a et $a^\dagger$ où tous les opérateurs d'annihilation sont à droite est dit avoir une forme *normale* ; dans la situation opposée où ce sont tous les $a^\dagger$ qui sont à droite, on parle de forme *anti-normale*. Tout produit d'opérateurs sous forme normale agissant sur le vide donne zéro.

(d) On définit les trois fonctions symétrique, normale et anti-normale :

$$C_\mathrm{S}(\lambda) \stackrel{\text{déf}}{=} \mathrm{Tr}\,\rho\,\mathrm{e}^{\lambda a^\dagger - \lambda^* a}\,,\; C_\mathrm{N}(\lambda) \stackrel{\text{déf}}{=} \mathrm{Tr}\,\rho\,\mathrm{e}^{\lambda a^\dagger}\,\mathrm{e}^{-\lambda^* a}\,,\; C_\mathrm{A}(\lambda) \stackrel{\text{déf}}{=} \mathrm{Tr}\,\rho\,\mathrm{e}^{-\lambda^* a}\,\mathrm{e}^{\lambda a^\dagger}\;. \tag{25.192}$$

En désignant toujours par x et p les variables dynamiques de l'oscillateur matériel associé à a et $a^\dagger$, mettre $C_\mathrm{S}(\lambda)$ sous la forme $\langle \mathrm{e}^{-\mathrm{i}(kx+\xi p)}\rangle \stackrel{\text{déf}}{=} \Gamma_\mathrm{S}(k,\xi)$. Montrer que la transformée de Fourier de $\Gamma_\mathrm{S}(k,\xi)$ est une fonction de Wigner (voir chapitre 21, section 21.4), $W(x,p) \equiv \mathcal{W}(\alpha)$.

(e) Montrer que, à un facteur près, les fonctions $C_\mathrm{N}(\lambda)$ et $C_\mathrm{A}(\lambda)$ sont respectivement les transformées de Fourier de $P(\alpha)$ et de $\langle\alpha|\rho|\alpha\rangle$.

(f) Montrer que la fonction $C_\mathrm{A}(\lambda)$ est proportionelle à $C_\mathrm{N}(\lambda)$; par le théorème de convolution, en déduire la relation entre $\langle\alpha|\rho|\alpha\rangle$ et $P(\alpha')$. Montrer que $\mathcal{W}(\alpha)$ et $P(\alpha')$ satisfont une relation du même type.

(g) Trouver $\langle\alpha|\rho|\alpha\rangle$ pour le rayonnement thermique ; en déduire $C_\mathrm{A}(\lambda)$, $C_\mathrm{N}(\lambda)$ et $P(\alpha)$ pour le champ à l'équilibre.

5. On suppose maintenant le champ couplé, à $t \geq 0$, à des sources *classiques* entièrement décrites par le courant[75] $\vec{J}(\vec{r},t)$. Le Hamiltonien s'écrit :

$$H = \sum_{\vec{k}\vec{\varepsilon}} \hbar\omega_k\, a^\dagger_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}} - \int_{\mathbb{R}^3} \vec{J}(\vec{r},t).\vec{A}(\vec{r})\,\mathrm{d}^3r\;. \tag{25.193}$$

(a) Mettre H sous la forme :

$$H = \hbar \sum_{\vec{k}\vec{\varepsilon}} \left[\omega_k\, a^\dagger_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}} - \Omega_{\vec{k}\vec{\varepsilon}}(t)\, a^\dagger_{\vec{k}\vec{\varepsilon}} - \Omega^*_{\vec{k}\vec{\varepsilon}}(t)\, a_{\vec{k}\vec{\varepsilon}}\right]\;. \tag{25.194}$$

(b) Écrire les équations de Heisenberg pour les opérateurs de création et d'annihilation, et les intégrer.

(c) En déduire que si le champ part de l'état vide $|\Psi(0)\rangle = |\mathrm{vac}\rangle$, il se trouve à l'instant t dans l'état cohérent $\otimes_{\vec{k}\vec{\varepsilon}}|\alpha_{\vec{k}\vec{\varepsilon}}(t)\rangle$, où $\alpha_{\vec{k}\vec{\varepsilon}}(t) = \mathrm{i}\int_0^t \Omega_{\vec{k}\vec{\varepsilon}}(t')\,\mathrm{d}t'$.

25.5.4 Calcul explicite de l'amplitude (25.168)

Il s'agit de calculer au second ordre l'amplitude donnée par l'égalité (25.168), avec l'opérateur de perturbation défini en (25.167). Comme le calcul est un peu long, on en détaille ci-dessous les différentes étapes ; les notations sont celles de la section 25.4.

1. Montrer que :

$$\langle \mathrm{g_A e}_{v\mathrm{B}}|V_\mathrm{I}(t_1)V_\mathrm{I}(t_2)|\mathrm{e}_{u\mathrm{A}}\mathrm{g_B}\rangle = \sum_{\vec{k}\vec{\varepsilon}} \frac{\hbar\omega_{\vec{k}}}{2\varepsilon_0 L^3} d_u d_v (\vec{e}_u.\vec{\varepsilon})(\vec{e}_v.\vec{\varepsilon})\,\mathrm{e}^{-\mathrm{i}\omega_{\vec{k}}(t_1-t_2)} \times \left[\mathrm{e}^{\mathrm{i}\omega_\mathrm{a}(t_1-t_2)}\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} + \mathrm{e}^{-\mathrm{i}\omega_\mathrm{a}(t_1-t_2)}\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}}\right] \tag{25.195}$$

[75] On est dans la situation "duale" du traitement semi-classique effectué dans la sous-section 24.5.1 ; ici, c'est le champ qui est quantifié et la matière qui ne l'est pas.

2. La correction au second ordre implique l'intégrale $\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 ...$ de l'expression précédente. En utilisant :

$$\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\, f(t_1, t_2) = \int_0^t \mathrm{d}t_2 \int_{t_2}^t \mathrm{d}t_1\, f(t_1, t_2) = \int_0^t \mathrm{d}t_1 \int_{t_1}^t \mathrm{d}t_2\, f(t_2, t_1) \tag{25.196}$$

   et en exploitant le fait que le vecteur $\vec{k}$ peut être changé en $-\vec{k}$ (juste là où il faut) dans la sommation $\sum_{\vec{k}\vec{\varepsilon}}$, montrer que l'on voit apparaître la quantité :

$$\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2\, \mathrm{e}^{\mathrm{i}k_\mathrm{a}c(t_1-t_2)}\, \mathrm{e}^{-\mathrm{i}kc|t_1-t_2|} \tag{25.197}$$

3. Afin de manipuler commodément la valeur absolue, on la représente comme suit[76] :

$$\mathrm{e}^{-\mathrm{i}kc|t_1-t_2|} = \mathrm{i}\frac{k}{\pi} \lim_{\varepsilon \to 0_+} \int_{-\infty}^{+\infty} \frac{\mathrm{e}^{-\mathrm{i}Kc(t_1-t_2)}}{K^2 - k^2 + \mathrm{i}\varepsilon} \mathrm{d}K \; ; \tag{25.198}$$

   ceci fait, effectuer les intégrales sur les temps.

4. L'égalité (25.153) est vraie quels que soient les vecteurs $\vec{D}$ et $\vec{D}\,'$: que devient-elle si on remplace ceux-ci par $\vec{e}_u$ et $\vec{e}_v$?

5. En déduire que l'amplitude cherchée est, au second ordre, donnée par :

$$A_{uv} \simeq A_{uv}^{(2)} = \frac{2\mathrm{i}\alpha}{\pi} \mathrm{e}^{-\mathrm{i}\omega_\mathrm{a}t} \lim_{\varepsilon \to 0_+} \int_{-\infty}^{+\infty} \frac{\sin^2 \frac{k_\mathrm{a}-K}{2}ct}{(k_\mathrm{a} - K)^2} r_u r_v\, g_{uv}(K, R)\, \mathrm{d}K \tag{25.199}$$

   où :

$$g_{uv}(K, R) \stackrel{\text{déf}}{=} \frac{1}{2\pi^2} \int (k^2 \delta_{uv} - k_u k_v) \frac{\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}}{k^2 - K^2 - \mathrm{i}\varepsilon} \mathrm{d}^3k \tag{25.200}$$

6. En remarquant que $k_u k_v \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} = -\partial R_u \partial R_v \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}$, établir l'expression (25.170) de la fonction $g_{uv}(K, R)$. Examiner, et interpréter, ses comportements suivant que $|K|R$ est très petit ou très grand devant 1.

7. Physiquement, et afin que la probabilité soit importante de préparer l'état où l'atome A est excité, l'excitation doit durer longtemps par rapport à $\omega_\mathrm{a}^{-1} \equiv (k_\mathrm{a}c)^{-1}$; cela étant, le noyau apparaissant dans l'intégrale sur K dans (25.199) est très fine. Par ailleurs, au voisinage de k_a, le facteur $\mathrm{e}^{\mathrm{i}|K|R}$ varie vite puisque $a_0 \ll R \ll \frac{1}{k_\mathrm{a}}$; au contraire, la fonction h_{uv} définie en (25.171) est un polynôme en $|K|R$ variant lentement, que l'on peut donc extraire de l'intégrale en le prenant à la valeur k_a. Enfin, oubliant le module de K (quand K est négatif, l'intégrand est très petit), on doit finalement calculer l'intégrale :

$$I \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} \frac{\sin^2 \frac{k_\mathrm{a}-K}{2}ct}{(k_\mathrm{a} - K)^2}\, \mathrm{e}^{\mathrm{i}KR}\, \mathrm{d}K \; ; \tag{25.201}$$

   montrer que $I = \frac{\pi}{2}c\,\mathrm{e}^{\mathrm{i}k_\mathrm{a}R}\big(t - \frac{R}{c}\big)\theta\big(t - \frac{R}{c}\big)$ (utiliser le théorème des résidus).

8. Des résultats précédents, déduire l'expression (25.169) de l'amplitude.

[76] Cette représentation intégrale se démontre avec le théorème des résidus ; l'établir à titre d'exercice.

Chapitre 26

Introduction à la théorie de la diffusion

Il s'agit ici de présenter les bases de la diffusion élastique de particules par une cible modélisée par un potentiel. Après avoir établi l'équation intégrale de la diffusion, la méthode des déphasages est exposée avant de définir l'approximation fondamentale de Born. La notion de résonance à basse énergie est reconsidérée et son scénario typique est exposé de façon élémentaire.

26.1 Présentation

L'importance des expériences de diffusion[1] de particules pour l'étude de la matière est universellement connue : c'est en faisant interagir une *pluie* de particules (faisceau incident) sur une cible et en analysant les caractéristiques du faisceau diffusé que l'on peut espérer obtenir des informations sur les interactions fondamentales, sur la dynamique interne de la cible, sur sa structure, etc. Le faisceau peut être de la lumière (photons) de fréquence *a priori* quelconque, mais évidemment adaptée au but poursuivi (rayons γ, rayons X, UV, etc.), soit des particules massives (neutrons, électrons, atomes), dont on ajuste la longueur d'onde associée de de Broglie pour la rendre commensurable avec la dimension typique des diffuseurs microscopiques (noyaux, atomes, ...) – c'est pourquoi les rayons X et les neutrons *thermiques* sont utilisés pour sonder la matière de basse énergie (voir chapitre 29, sous-section 29.3.2). La figure 26.1 donne le schéma de principe d'une expérience de diffusion.

[1]On parlera indifféremment de théorie de la diffusion ou de théorie des collisions. Le mot *collision* est évidemment à prendre au sens usuel en Physique, et non dans l'acception restrictive d'un *choc* (voir

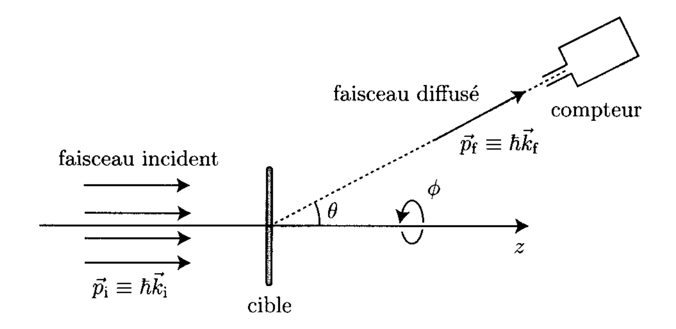


Figure 26.1: Schéma d'une expérience de diffusion. Une diffusion élastique est caractérisée par l'égalité $\|\vec{p}_f\| = \|\vec{p}_i\|$

Les phénomènes de diffusion sont en général fort complexes ; notamment, les particules du faisceau peuvent être composites et se fragmenter sous l'effet de la collision – tout comme les constituants de la cible, d'ailleurs. Afin de rester sur un plan introductif et de se borner aux idées premières, on adopte pour la suite les hypothèses simplificatrices suivantes :

1. On ignore le spin des particules. Ceci n'est pas toujours possible : par exemple, la diffusion des neutrons, outre une interaction de contact à l'échelle atomique[2], procède également d'une interaction entre le spin du neutron et les spins nucléaires ou électroniques (voir chapitre 29, sous-section 29.3.2). Dans ce dernier cas, la diffusion correspondante donne des informations précieuses sur les fonctions d'onde des électrons (extension, anisotropie, etc.). De surcroît, l'usage d'un faisceau polarisé en spin permet de sonder les structures magnétiques de la matière condensée, ou ses excitations (magnons[3], par exemple).

2. Toutes les particules en jeu (projectiles et constituants de la cible) sont supposées sans structure interne : seules les collisions *élastiques* seront donc considérées, où par définition, l'énergie cinétique des projectiles est la même avant et après la collision (on néglige ainsi tout effet de *recul*). Physiquement, ceci signifie que les éventuels degrés de liberté internes des partenaires en présence ne peuvent être excités qu'avec des énergies très supérieures à celles mises en jeu lors de la collision, ou sont inaccessibles en raison de règles de sélection spécifiques.

3. Les diffuseurs de la cible sont supposés infiniment massifs (et sans structure, comme déjà admis). Ceci permet de remplacer chacun d'entre eux par le champ qu'il crée, donnant un potentiel[4] $V(\vec{r})$. Cette hypothèse n'est pas restrictive : quand il n'existe que des interactions à deux corps, on peut toujours se placer dans le repère du

les remarques à ce sujet dans le Tome I, p. 56).

[2]Cette interaction a une portée de l'ordre du rayon nucléaire et peut de ce fait être considérée comme ponctuelle à l'échelle de l'angström.

[3]Ces excitations magnétiques sont discutées dans le chapitre 30, section 30.6.1.

[4]plus précisément, une énergie potentielle.

centre de masse où seule compte la dynamique de la particule fictive ayant la masse réduite μ (voir Tome I, éq. (3.19)). Pour revenir aux observations effectuées dans le repère dit *du laboratoire*, il suffit d'utiliser la relation de passage entre la section efficace calculée dans le repère du centre de masse et celle qui est mesurée (voir Tome I, problème 3.6.7).

4. La cible est supposée assez mince pour que chaque projectile ne subisse qu'une collision ; cette situation est communément réalisée en pratique puisque, quand l'interaction projectile - cible est assez faible, l'obtention d'un signal exploitable nécessite un courant de sortie relativement important.

5. On néglige toute corrélation entre les particules diffusées, ce qui exige notamment que le faisceau soit de faible densité[5]. Cette hypothèse est presque une nécessité méthodologique : le plus souvent, on veut observer la *signature* de la cible pour pouvoir obtenir des informations propres à celle-ci, ce qui exige que le faisceau de projectiles constitue une *petite* perturbation.

Mutatis mutandis, la plupart des notions introduites ci-dessous conservent leur pertinence au dehors du cadre délibérément restrictif ainsi précisé.

Les problèmes de diffusion peuvent être abordés de deux façons, essentiellement. La première consiste à considérer exclusivement des états *stationnaires* (approche statique), ce qui veut dire physiquement que le faisceau incident a été allumé depuis assez longtemps pour que le système global se trouve en régime permanent, tous transitoires éteints. On est alors réduit à résoudre une équation aux valeurs propres, donnant les états stationnaires du système faisceau - cible. La deuxième approche, dynamique, est plus physique : elle consiste à représenter l'interaction particule - cible par une perturbation *dépendant du temps* (pour un traitement à une dimension, voir [206]), simulant le passage d'une particule venue de l'infini et y repartant après collision. Cette deuxième façon de faire est lourde techiquement, et sera omise ici ; son développement peut être trouvé dans l'ouvrage classique de Roman [207]. Bien évidemment, les deux approches se rejoignent quand, partant de la première, on introduit des paquets d'ondes construits avec les états stationnaires, représentant une particule émise par la source se dirigeant vers la cible, avant de s'en éloigner pour être détectée *à l'infini*.

26.2 Diffusion par un potentiel

Compte tenu des hypothèses énoncées ci-dessus, l'approche statique consiste à trouver les états quantiques d'un projectile dans le champ d'un diffuseur désincarné physiquement et représenté par le champ de force conservatif dont il est la source, donnant à un projectile

[5]La faible densité (schématiquement : le non-recouvrement spatial des fonctions d'onde des projectiles) autorise également à délaisser toute complication liée à l'indiscernabilité des projectiles.

de masse m l'énergie potentielle $V(\vec{r})$ (*potential scattering*). On est donc en situation de devoir résoudre l'équation aux valeurs propres pour un certain Hamiltonien H :

$$\boxed{H = \frac{\vec{p}^{\,2}}{2m} + V(\vec{r})} \tag{26.1}$$

L'un des enjeux de la théorie des collisions est l'obtention de la section efficace différentielle de diffusion, $\sigma_\mathrm{d}(\theta,\, \phi)$, quantité dont la définition a été donnée dans le Tome I (section 3.4), dans un cadre classique. Cette définition, de nature statistique, n'est nullement tributaire de l'aspect classique ou quantique des particules et peut donc être ici conservée ; bien évidemment, l'aspect quantique des projectiles est essentiel et joue à plein quand leur longueur d'onde de de Broglie est commensurable avec une longueur caractéristique a du potentiel diffuseur qui, en pratique, définira sa *portée* spatiale. La longueur a étant fixée, cette adéquation est acquise en ajustant l'énergie E des projectiles de sorte que $\frac{h}{\sqrt{2mE}} \sim a$.

Notant ici $\mathrm{d}n(\theta,\, \phi)$ le nombre de projectiles diffusés par unité de temps dans l'angle solide $\mathrm{d}\Omega$ déterminé par les angles θ (déviation) et ϕ (azimut), et j_source le courant incident (nombre de projectiles traversant l'unité de section droite du faisceau par unité de temps), $\sigma_\mathrm{d}(\theta,\, \phi)$ satisfait par définition :

$$\boxed{\mathrm{d}n(\theta,\, \phi) = j_\mathrm{source}\, \sigma_\mathrm{d}(\theta,\, \phi)\, \mathrm{d}\Omega} \tag{26.2}$$

σ_d est homogène à une surface. La section efficace totale de collision s'obtient par intégration sur les angles :

$$\sigma = \int \sigma_\mathrm{d}(\theta,\, \phi)\, \mathrm{d}\Omega = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin\theta\, \mathrm{d}\theta\, \sigma_\mathrm{d}(\theta,\, \phi) \quad . \tag{26.3}$$

De toute évidence, σ donne une information globale, alors que $\sigma_\mathrm{d}(\theta,\, \phi)$ fournit des renseignements très fins, notamment sur l'éventuelle anisotropie de la cible[6].

Le potentiel $V(\vec{r})$ est supposé (par simplicité) tendre vers la *même* valeur dans toutes les directions, valeur que l'on prend conventionnellement égale à zéro. En conséquence, les états stationnaires de diffusion ne sont rien d'autre que les états non-liés du potentiel ; il s'agit donc des états propres $\psi(\vec{r})$ du Hamiltonien (26.1), tous associés à une certaine valeur propre E, énergie qui est *positive* ; en représentation-q, l'équation à résoudre est :

$$\left[\frac{\vec{p}^{\,2}}{2m} + V(\vec{r})\right] \psi(\vec{r}) = E\psi(\vec{r}) \quad , \tag{26.4}$$

où $\vec{p} = -\mathrm{i}\hbar\vec{\nabla}$; E est l'énergie d'un projectile, qui est la même avant et après la diffusion puisque celle-ci est réputée élastique. En posant[7] :

$$E \stackrel{\text{déf}}{=} \frac{\hbar^2 k^2}{2m} \qquad (k \in \mathbb{R}) \ , \qquad V(\vec{r}) \stackrel{\text{déf}}{=} g\frac{\hbar^2}{2m}U(\vec{r}) \ , \tag{26.5}$$

[6]On verra que pour une cible à symétrie sphérique ($V(\vec{r}) \equiv V(r)$), σ_d ne dépend pas de l'angle ϕ.

[7]Comme dans le chapitre 24, la constante de couplage sans dimension g est introduite pour la commodité de certains raisonnements. On pourra toujours faire formellement $g = 1$ dans les formules obtenues.

l'équation (26.4) se récrit comme suit :

$$\boxed{\left[\Delta + \vec{k}^2 - gU(\vec{r})\right] \psi(\vec{r}) = 0} \qquad (26.6)$$

On sait résoudre exactement cette équation pour quelques potentiels remarquables (voir par exemple sous-section 19.1.4 le cas du puits "carré"), les fonctions propres exactes différant bien évidemment d'un potentiel à l'autre. Ces différences sont de peu d'intérêt, puisque le compteur se trouve toujours, en pratique, à une distance gigantesque comparée à l'échelle caractéristique de variation de l'interaction due à la cible et décrite par $V(\vec{r})$. En d'autres termes, ce qui importe ce sont les propriétés *asymptotiques* ("à l'infini") des solutions propres $\psi(\vec{r})$, dont on peut deviner la forme générale sur la base de considérations physiques simples, *indépendantes* de la nature précise du potentiel et, de ce fait, universelles – avec toutefois des précautions à prendre pour des potentiels à décroissance lente, comme le champ Coulombien. Dans ce qui suit, on raisonne avec l'axe Oz parallèle au faisceau incident et passant par le centre diffuseur (voir fig. 26.2).

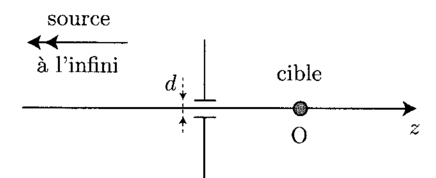


Figure 26.2: Géométrie utilisée dans le texte. La source est située à l'infini en $z = -\infty$.

Examinons d'abord le problème en l'absence de potentiel diffuseur ($g = 0$, strictement). Les solutions de (26.6) sont des ondes planes de la forme $A\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} + B\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}$, où A et B sont des constantes pour l'instant quelconques. Avec la géométrie utilisée (source en $z = -\infty$), on a bien sûr $B = 0$, d'une part ; d'autre part le vecteur $\vec{k}$ est le long de Oz : la seule solution est donc $A\,\mathrm{e}^{\mathrm{i}kz}$ ($k > 0$). La constante A est fixée par la valeur du courant[8] émis par la source, j_{source}, soit

$$j_{\text{source}} = \frac{\hbar k}{m}|A|^2 \ , \qquad (26.8)$$

d'où la constante A, à une phase inessentielle près.

Supposons maintenant que la constante de couplage g est non-nulle, mais *infiniment petite* : $g = 0_+$. Dans ces conditions, la symétrie de translation continue est brisée, puisqu'il y a maintenant un point remarquable (là où se trouve le centre diffuseur, pris comme origine dans l'espace). La solution change donc *qualitativement* d'allure : outre l'onde plane $\propto \mathrm{e}^{\mathrm{i}kz}$, doit apparaître – avec une amplitude infiniment petite puisque

[8]Rappelons au passage que l'expression du vecteur courant est :

$$\vec{j} = \frac{\hbar}{2\mathrm{i}m}\left[\psi^*(\vec{r})\vec{\nabla}\psi(\vec{r}) - \psi(\vec{r})\vec{\nabla}\psi^*(\vec{r})\right] \equiv \frac{\hbar}{m}\,\Im\left[\psi^*(\vec{r})\vec{\nabla}\psi(\vec{r})\right] \ . \qquad (26.7)$$

$g = 0_+$ – une onde *sphérique* issue du point $\vec{r} = \vec{0}$. La forme générale de la solution est alors :

$$g = 0_+ \ : \quad \psi_{\vec{k}}(\vec{r}) = A\,\mathrm{e}^{\mathrm{i}kz} + C_k(\vec{r})\,\mathrm{e}^{\mathrm{i}kr} \ , \tag{26.9}$$

où l'amplitude $C_k(\vec{r})$ est infiniment petite et, évidemment, décroît avec $r \equiv \|\vec{r}\,\|$.

Comme expliqué ci-dessus, ce sont les propriétés asymptotiques des solutions de (26.6) qui importent. Supposant maintenant g *fini*, on doit s'attendre à ce que, pourvu que le potentiel $U(\vec{r})$ décroisse assez vite à l'infini[9], la forme (26.9) soit toujours vraie mais cette fois en tant qu'expression asymptotique de la solution exacte à g fini. Ce point sera vérifié en détail dans la suite, mais un argument physique permet déjà de prévoir le comportement de la fonction $C(\vec{r})$ en fonction de r : de toute évidence, la conservation du courant impose $|C(\vec{r})| \propto \frac{1}{r}$ quand $r \to +\infty$; c'est la seule dépendance en r qui assure que le flux du courant à l'infini à travers une sphère de très grand rayon est une constante, indépendante du rayon de la sphère[10]. À ce stade, on peut ainsi anticiper sur des bases physiques que la solution exacte a l'expression asymptotique suivante :

$$\boxed{r \to +\infty \ : \ \psi_{\vec{k}}(\vec{r}) \sim A\Big[\mathrm{e}^{\mathrm{i}kz} + f_k(\theta,\,\phi)\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\Big]} \tag{26.10}$$

où la fonction $f_k(\theta,\,\phi)$ est indépendante de r et caractérise tout autant l'anisotropie du potentiel diffuseur (par l'angle ϕ) que la variation du taux de comptage en fonction de l'angle de déviation θ.

Montrons techniquement qu'il en est bien ainsi. En présence du potentiel, le report de la forme asymptotique (26.10) dans le premier membre de (26.6) donne :

$$\big[\Delta + k^2 - gU(\vec{r})\big]\,\big[\mathrm{e}^{\mathrm{i}kz} + f_k(\theta,\,\phi)\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\big] \equiv \big[\frac{2}{r^2} - 2\frac{\mathrm{i}k}{r} - gU(\vec{r})\big]\,\frac{1}{r}\mathrm{e}^{\mathrm{i}kr} \tag{26.11}$$

Supposant que $U(\vec{r}) \sim U_\infty\big(\frac{a}{r}\big)^\alpha$ quand $r \gg a$, on voit que cette solution asymptotique est correcte pourvu que $r^{-(\alpha-1)}$ tende vers zéro, ce qui impose $\alpha > 1$. Cette condition exclut donc le potentiel Coulombien nu, qui doit faire l'objet d'une attention particulière (voir la brève discussion p. 1213) ; cette restriction n'est toutefois pas très sévère car, en pratique, l'interaction de Coulomb est le plus souvent affectée par l'écrantage.

Ainsi, avec un potentiel tendant assez vite vers zéro à l'infini, et pour une source située à l'infini (à gauche), la forme asymptotique de toute solution est donnée par

[9] Cette hypothèse place le potentiel Coulombien *nu* à part, en raison de sa portée *infinie*. On verra en fait que, en dépit de cet aspect spécifique qui ruine *a priori* les arguments généraux développés ci-dessous, il reste vrai que, même dans ce cas, la section efficace est donnée par l'égalité fondamentale (26.16) (voir la brève discussion dans la Remarque p. 1213). Pour la discussion en cours, signalons juste que l'obligation $|C(r)| \propto r^{-1}$ est une absolue nécessité, compatible avec le fait que ce régime asymptotique est parfois réalisé seulement très loin en raison des spécificités à longue distance du potentiel.

[10] Le facteur r^{-1} provient des trois dimensions de l'espace. Dans $\mathbb{R}^D$, l'onde sphérique asymptotique est $r^{-(D-1)/2}\,\mathrm{e}^{\mathrm{i}kr}$ (voir l'expression de $\Delta \equiv -2m\hbar^{-2}p_r^2$ dans le chapitre 19, éq. (19.21)). C'est l'exposant $(D-1)/2$ – dans $\mathbb{R}^D$, $\mathrm{d}^D r \propto r^{D-1}\mathrm{d}r$ – qui assure que le flux du courant radial à travers la grande sphère de rayon R est indépendant de R, propriété exprimant la conservation du nombre de projectiles quelle que soit la distance où on les compte tous, en particulier à l'infini.

Le point important est que le courant est *quadratique* en $\psi(\vec{r})$ (voir note 8).

l'expression (26.10), avec $k > 0$. La dépendance en r^{-1} confirme l'argument fondé sur la conservation du courant. Pour $r \to +\infty$, le terme dominant est $\mathrm{e}^{\mathrm{i}kz}$; lorsque z est négatif, c'est juste le faisceau incident ; pour $z > 0$, ce terme représente les particules transmises sans modification[11] (diffusion vers l'avant). La composante sphérique $\frac{1}{r}\,\mathrm{e}^{\mathrm{i}kr}$, à elle seule, représente l'onde effectivement diffusée.

Le coefficient f_k (homogène à une longueur) dépend des angles en général et c'est lui (et lui seul) qui colporte l'information sur la cible agissant par l'intermédiaire du potentiel (f_k constitue l'*image brute* de la cible). f_k s'appelle *amplitude de diffusion* ; l'indice k rappelle que c'est une fonction du vecteur d'onde, donc de l'énergie. Finalement, l'expression (26.10) est une combinaison linéaire de deux solutions (de même énergie) : c'est bien la solution (asymptotique) de l'équation aux valeurs propres (26.4), compte tenu de la position de la source (laquelle constitue *ipso facto* une condition aux limites). Pour des raisons évidentes, la composante sphérique ayant le comportement asymptotique (26.10) est appelée *onde sortante*, et est valide *au loin*, à une distance r très grande devant la longueur caractéristique a du potentiel.

En pratique, l'onde plane $\mathrm{e}^{\mathrm{i}kz}$ n'est pas infiniment étendue dans la direction transverse, le faisceau incident ayant une ouverture finie (fixée par la longueur d, voir fig. 26.2), que l'on retrouve dans l'onde plane transmise vers $z > 0$. Afin de ne pas éblouir le détecteur par un signal sans intérêt, celui-ci est placé en-dehors de la direction d'incidence. Dans toute la suite, on ne compte donc que les particules dont l'angle de diffusion θ est plus grand que $\theta_{\mathrm{min}} \stackrel{\text{déf}}{=} \mathrm{Arcsin}(\frac{d}{2D})$, où D est la distance entre le compteur et le potentiel. L'angle exclu dépend des conditions expérimentales précises, mais peut manifestement, en principe, être rendu aussi petit que l'on veut.

Il existe une relation remarquable très simple entre l'amplitude $f_k(\theta,\,\phi)$ et la section efficace différentielle $\sigma_{\mathrm{d}}(\theta,\,\phi)$. En effet, si on se place en $z = -\infty$ ($1/r = 0$), seule subsiste l'onde incidente $A\mathrm{e}^{\mathrm{i}kz}$, donnant le courant parallèle à Oz de module j_{source}, relié à la constante A suivant l'égalité (26.8). En ce qui concerne le courant diffusé, on ne doit pas *a priori* omettre les interférences entre l'onde plane et l'onde sphérique. Toutefois, en excluant les directions de diffusion vers l'avant ($|\theta| < \theta_{\mathrm{min}}$) et vers l'arrière ($|\pi - \theta| < \theta_{\mathrm{min}}$), l'intégralité du courant diffusé est véhiculé par l'onde sphérique, à des quantités près qui sont aussi petites que l'on veut. Les composantes du gradient étant :

$$\vec{\nabla} = \left(\frac{\partial}{\partial r},\, \frac{1}{r}\frac{\partial}{\partial\theta},\, \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right) \ , \tag{26.12}$$

le courant diffusé porté par l'onde sphérique $Af_k(\theta,\,\phi)\,r^{-1}\,\mathrm{e}^{\mathrm{i}kr}$ est :

$$\vec{j}_{\mathrm{d}} = |A|^2\Big[\frac{\hbar k}{mr^2}|f_k|^2,\, \frac{\hbar}{mr^3}\,\Im\left(f_k^*\frac{\partial f_k}{\partial\theta}\right),\, \frac{\hbar}{mr^3\sin\theta}\,\Im\left(f_k^*\frac{\partial f_k}{\partial\phi}\right)\Big] \ . \tag{26.13}$$

Dans la zone de comptage, seule subsiste en pratique la composante radiale, les deux autres étant beaucoup plus petites du fait de leur dépendance en r^{-3}. On a donc :

$$\vec{j}_{\mathrm{d}} \simeq \vec{j}_{\mathrm{radial}} = |A|^2\frac{\hbar k}{mr^2}\,|f_k(\theta,\,\phi)|^2\,\frac{\vec{r}}{r} \ . \tag{26.14}$$

[11] Noter que l'on ne peut pas dissocier les deux composantes : il y a donc toujours une onde transmise sans altération.

Le nombre des particules par unité de temps, dn, qui frappent le détecteur orienté suivant la direction $(\theta,\, \phi)$, est $j_{\mathrm{d}}\,\mathrm{d}S$, où $\mathrm{d}S = r^2\mathrm{d}\Omega$ est l'ouverture de celui-ci[12]. Compte tenu de (26.8), il vient :

$$\mathrm{d}n = j_{\mathrm{d}}\, r^2\mathrm{d}\Omega = |A|^2\frac{\hbar k}{mr^2}\,|f_k(\theta,\, \phi)|^2\, r^2\mathrm{d}\Omega \equiv j_{\mathrm{source}}|f_k(\theta,\, \phi)|^2\,\mathrm{d}\Omega\ , \tag{26.15}$$

d'où, en se référant à (26.2), la relation simple et remarquable entre amplitude de diffusion et section efficace différentielle :

$$\boxed{\sigma_{\mathrm{d}}(\theta,\, \phi) = |f_k(\theta,\, \phi)|^2} \tag{26.16}$$

Ainsi, le module carré de l'amplitude de l'onde sphérique diffusée n'est rien d'autre que la section efficace cherchée : l'enjeu est bel et bien de trouver $f_k(\theta,\, \phi)$. Ceci permettra de comparer $|f_k(\theta,\, \phi)|^2$ à la section efficace mesurée et de valider un modèle ou un autre.

En définitive, la détermination de la section efficace est maintenant réduite à trouver, non pas la solution complète de l'équation aux fonctions propres, mais seulement son comportement asymptotique, pour une énergie E donnée ; de ce point de vue, la situation est très différente de celle rencontrée à propos de la détermination des états propres (liés ou non-liés) d'un potentiel $V(\vec{r}\,)$, où la toute première question est de trouver les énergies propres (avant d'en déduire les fonctions propres). Ici, on connaît d'avance la *forme* asymptotique de la solution, et l'énergie est donnée (c'est celle des particules au moment de leur émission par la source) ; pour marquer ce fait, on note désormais E_0 l'énergie initiale des particules. Dans ces conditions, il n'est pas surprenant de constater que tous les aspects physiques liés au potentiel sont exclusivement contenus dans l'amplitude de diffusion $f_{k_0}(\theta,\, \phi)$ $(k_0 = \hbar^{-1}\sqrt{2mE_0}\,)$.

♦ *Remarque*

La description de la diffusion en termes d'états stationnaires est clairement une idéalisation de la réalité. Physiquement, il convient de former des paquets d'ondes[13] ; pour la pertinence de l'analyse avec des états stationnaires, il faut au moins que toutes les longueurs importantes physiquement soient petites devant celles qui caractérisent le paquet d'ondes. Si l et d sont les dimensions longitudinale et transversale de celui-ci, il faut certainement que la portée a du potentiel satisfasse $a \ll l,\, d$, afin que l'aspect *localisé* du paquet d'ondes soit, pour le potentiel, indiscernable d'une onde plane infiniment étendue. Une analyse détaillée [34] montre que la relation fondamentale (26.16) reste vraie pourvu que l'on ait en plus :

$$\boxed{\lambda D \ll l^2,\, d^2} \tag{26.17}$$

[12]puisque dΩ est par définition la surface de sphère de rayon unité délimitée par le petit cône d'axe $(\theta,\, \phi)$.

Par ailleurs, on se place toujours en dehors des directions avant et arrière, défini par un angle aussi voisin que l'on veut de 0 et π. Le résultat ci-dessus exprime la conservation du courant : j_{source} est égal au courant sphérique, une fois éliminées les deux contributions des deux directions proches de 0 et π, qui sont aussi petites que l'on veut.

[13]À ce sujet, voir également la discussion physique dans [4], p. 219, et l'article de Katz [208].

où $\lambda \stackrel{\text{déf}}{=} \frac{h}{\sqrt{2mE_0}}$ est la longueur d'onde de de Broglie d'un projectile, et avec toujours $d \ll D \sin\theta$. Cette dernière condition n'est visiblement pas un problème, et exclut juste un tout petit angle solide le long du faisceau incident, de part et d'autre de la cible. D'un autre côté, s'agissant de *photographier* à 1 m de distance des cibles dont la longueur typique est de l'ordre de 1 Å au plus, la condition (26.17) exige $d,\, l \gg 10^{-5}$ m : la taille typique des paquets d'ondes doit donc être au moins de l'ordre de quelques dizaines de microns. À cette taille typique Δx correspond une incertitude en impulsion, Δp, donc une incertitude en énergie ΔE ; l'incertitude relative $\frac{\Delta E}{E}$ est $\sim \frac{\lambda}{\Delta x} \lesssim 10^{-5}$. ♦

26.3 Équation intégrale de la diffusion

Avec $E_0 = \frac{\hbar^2 k_0^2}{2m}$ $(k_0 > 0)$, l'équation (26.6) s'écrit aussi :

$$(\Delta + k_0^2)\, \psi_{\vec{k}_0}(\vec{r}) = gU(\vec{r})\psi_{\vec{k}_0}(\vec{r}) \ ; \tag{26.18}$$

pour des raisons qui seront plus claires dans la suite, la méthode de la fonction de Green se révèle particulièrement commode pour manipuler formellement cette équation[14]. L'idée de base de cette technique consiste à considérer le second membre de (26.18) comme une *source* ; dans ces conditions, la fonction de Green $G(\vec{r})$ étant solution de l'équation[15] :

$$(\Delta + k_0^2)\, G(\vec{r}) = \delta(\vec{r}) \ , \tag{26.19}$$

la solution générale de l'équation (26.18) admet l'expression formelle :

$$\boxed{\psi_{\vec{k}_0}(\vec{r}) = \psi_{\vec{k}_0}^{(0)}(\vec{r}) + g \int_{\mathbb{R}^3} G(\vec{r} - \vec{r}\,')U(\vec{r}\,')\psi_{\vec{k}_0}(\vec{r}\,')\, \mathrm{d}^3 r'} \tag{26.20}$$

où $\psi_{\vec{k}_0}^{(0)}(\vec{r})$ est la solution générale de l'équation homogène, satisfaisant par définition :

$$(\Delta + k_0^2)\, \psi_{\vec{k}_0}^{(0)}(\vec{r}) = 0 \ , \tag{26.21}$$

et des conditions aux limites reflétant la physique du problème considéré. Bien sûr, le problème n'est pas pour autant résolu, puisque l'inconnue $\psi_{k_0}(\vec{r})$ figure dans les deux membres de (26.20) ; néanmoins, on pressent que cette écriture (tout comme pour l'opérateur d'évolution, comparer avec (24.24)) va permettre de définir des approximations successives, dans l'hypothèse où $U(\vec{r}\,')$ est une perturbation non singulière.

Toutes les équations étant du second ordre, il convient de préciser des conditions aux limites assurant l'unicité de la solution. En ce qui concerne la fonction $\psi_{k_0}(\vec{r})$, on lui impose de posséder à l'infini le comportement asymptotique précisé en (26.10), dont on a

[14]La même technique a d'ailleurs déjà été utilisée dans la sous-section 23.2.4.

[15]La fonction $G(\vec{r})$ est paramétrée par k_0 mais cette dépendance est sous-entendue pour alléger les notations.

vu qu'il est universel pour des potentiels tendant plus vite vers zéro que r^{-1}. Quant à la fonction de Green, qui est indépendante du potentiel par définition, il convient de choisir celle qui, parmi toutes les solutions de (26.19), engendre précisément une solution $\psi_{k_0}(\vec{r})$ ayant ce comportement asymptotique. L'indépendance de G vis-à-vis du potentiel permet d'énoncer les conditions sur G pour obtenir la *bonne* fonction de Green pour le problème considéré en considérant un cas *particulier* de potentiel. Prenons $U(\vec{r}) \propto \delta(\vec{r})$; dans ces conditions, l'intégrale dans (26.20) est proportionnelle à $G(\vec{r})$. Comme $\psi_{k_0}(\vec{r})$ doit être de la forme (26.10), on voit que la bonne fonction de Green est ici proportionnelle à $e^{ik_0 r}$. La suite montrera que cette seule condition, nécessaire, est aussi suffisante.

Ce point étant précisé, il faut donc résoudre l'équation (26.19), dont la nature est telle que sa solution est à considérer *a priori* comme une *distribution* ; pour la résolution, la transformation de Fourier s'impose, incitant à introduire :

$$G(\vec{r}) \stackrel{\text{déf}}{=} \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \mathcal{G}(\vec{k}) e^{i\vec{k}.\vec{r}} \, d^3k \ . \tag{26.22}$$

La transformée de Fourier de l'équation (26.19) est :

$$(-k^2 + k_0^2)\mathcal{G}(\vec{k}) = 1 \ , \tag{26.23}$$

qui montre que $\mathcal{G}(\vec{k})$ ne dépend que du module $k \stackrel{\text{déf}}{=} \|\vec{k}\|$; une fois l'intégration angulaire effectuée[16], (26.22) prend la forme :

$$G(\vec{r}) \equiv G(r) = \frac{1}{2\pi^2 r} \int_0^{+\infty} \mathcal{G}(k) \sin kr \ k dk \ . \tag{26.24}$$

L'équation (26.23) a une solution *particulière* égale à la partie principale de Cauchy, $\mathcal{G}_{\text{part}}(k) = \mathcal{P}\frac{1}{k_0^2 - k^2}$; la solution *générale* s'obtient en ajoutant à cette solution particulière l'intégrale générale de l'équation sans second membre $(-k^2 + k_0^2)\mathcal{G}_{\text{hom}}(k) = 0$, qui est une combinaison linéaire arbitraire des fonctions de Dirac $\delta(k \pm k_0)$, puisque $x\delta(x) = 0$. La solution générale de (26.23) est ainsi :

$$\mathcal{G}(k) = \mathcal{P}\frac{1}{k_0^2 - k^2} + C_+\delta(k - k_0) + C_-\delta(k + k_0) \ , \tag{26.25}$$

avec les $C_\pm$ pour l'instant arbitraires. Selon (26.24), et avec toujours le choix conventionnel $k_0 > 0$, on en déduit :

$$G(r) = \frac{1}{2\pi^2 r}\left[\frac{1}{2i}\mathcal{P}\int_{-\infty}^{+\infty} \frac{k\, e^{ikr}}{k_0^2 - k^2} \, dk + C_+ k_0 \sin k_0 r\right] \ . \tag{26.26}$$

La partie principale se calcule aisément par résidus et vaut $-i\pi \cos k_0 r$, d'où l'expression générale exacte de $G(r)$:

$$G(r) = \frac{1}{2\pi^2 r}\left(- \frac{\pi}{2} \cos k_0 r + C_+ k_0 \sin k_0 r\right) \ . \tag{26.27}$$

[16]La partie angulaire de l'intégrale est $\int_0^\pi \sin\theta \, d\theta \int_0^{2\pi} d\phi \, e^{ikr\cos\theta} = \frac{4\pi}{kr} \sin kr$.

On sait de l'analyse préalable que la fonction $G(r)$ doit être asymptotiquement proportionnelle à $\mathrm{e}^{\mathrm{i}k_0 r}$, ce qui impose $C_+ = -\frac{\mathrm{i}\pi}{2k_0}$, d'où finalement la *bonne* fonction de Green $G^{(+)}(r)$ pour le problème de diffusion considéré[17] :

$$\boxed{G^{(+)}(r) = -\frac{1}{4\pi r}\,\mathrm{e}^{\mathrm{i}k_0 r}} \tag{26.29}$$

qui produit bien, comme il se doit, une onde *sortante.*

Revenant maintenant à (26.20), on obtient l'équation intégrale prenant en compte toutes les conditions aux limites imposées par la situation physique, et pour une énergie E_0 donnée[18] :

$$\boxed{\psi_{\vec{k}_0}(\vec{r}) = \mathrm{e}^{\mathrm{i}k_0 z} - g\,\frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{\mathrm{e}^{\mathrm{i}k_0\|\vec{r}-\vec{r}\,'\|}}{\|\vec{r}-\vec{r}\,'\|}\,U(\vec{r}\,')\psi_{\vec{k}_0}(\vec{r}\,')\,\mathrm{d}^3r'} \tag{26.30}$$

Répétons que l'inconnue figure dans les deux membres, de sorte qu'il reste toujours à trouver la solution. On peut toutefois, sur cette équation, examiner la limite asymptotique $r \to +\infty$, afin de retrouver dans ce cadre l'amplitude de diffusion $f_{k_0}(\theta,\, \phi)$. Dans cette limite, et toujours dans l'hypothèse où le potentiel décroît assez vite, $U(\vec{r}\,')$ joue un peu le rôle d'une fonction de coupure sur l'échelle a ; quand $r \gg a$, on peut écrire :

$$\|\vec{r}-\vec{r}\,'\| \simeq r - \vec{n}.\vec{r}\,' \ , \tag{26.31}$$

où $\vec{n} \stackrel{\text{déf}}{=} \frac{\vec{r}}{r}$ est le vecteur unitaire de la direction de comptage. La forme asymptotique de la solution exacte déduite de l'équation intégrale est alors :

$$r \gg a\ :\quad \psi_{\vec{k}_0}(\vec{r}) \sim \mathrm{e}^{\mathrm{i}k_0 z} - g\,\frac{1}{4\pi}\frac{\mathrm{e}^{\mathrm{i}k_0 r}}{r}\int_{\mathbb{R}^3} \mathrm{e}^{-\mathrm{i}k_0\vec{n}.\vec{r}\,'}\,U(\vec{r}\,')\psi_{\vec{k}_0}(\vec{r}\,')\,\mathrm{d}^3r' \ . \tag{26.32}$$

Par comparaison avec (26.10) (avec $A = 1$), on en déduit l'égalité :

$$\boxed{f_{k_0}(\theta,\, \phi) = -g\,\frac{1}{4\pi}\int_{\mathbb{R}^3} \mathrm{e}^{-\mathrm{i}k_0\vec{n}.\vec{r}\,'}\,U(\vec{r}\,')\psi_{\vec{k}_0}(\vec{r}\,')\,\mathrm{d}^3r' \equiv -g\,\frac{1}{4\pi}\langle\psi^{(0)}_{k_0\vec{n}}|U|\psi_{\vec{k}_0}\rangle} \tag{26.33}$$

À un facteur près, l'amplitude de diffusion est donc le produit scalaire de $U|\psi_{\vec{k}_0}\rangle$, où $|\psi_{\vec{k}_0}\rangle$ est la solution exacte d'énergie E_0, avec l'état *libre* de même énergie $|\psi^{(0)}_{k_0\vec{n}}\rangle$, détecté au niveau du compteur situé quasiment à l'infini dans la direction fixée par $\vec{n}$.

---

[17] Pour $k_0 = 0$, l'équation satisfaite par $G^{(+)}$ est, à des constantes près, l'équation de Poisson pour une charge ponctuelle. On retiendra au passage l'égalité utile :

$$\boxed{\Delta\frac{1}{r} = -4\pi\delta(\vec{r})} \tag{26.28}$$

Cette égalité est le cas particulier $\kappa = 0$ de (23.159). Par ailleurs, et par comparaison avec la sous-section 23.2.4 où la technique de la fonction de Green a déjà été utilisée, on voit qu'ici la géométrie précise de l'expérience modélisée ne laisse plus aucune constante indéterminée (comparer avec (23.163)).

[18] Afin de simplifier les écritures, on délaisse désormais toutes les constantes de proportionnalité, comme A dans (26.10), ce qui revient à poser $A = 1$ partout.

L'équation (26.30) est une équation intégrale de Fredholm de seconde espèce [184] ; en écrivant :

$$\psi_{\vec{k}_0}(\vec{r}) = \mathrm{e}^{\mathrm{i}k_0 z} - g\,\frac{1}{4\pi}\int_{\mathbb{R}^3} G^{(+)}(\vec{r}-\vec{r}_1)\psi_{\vec{k}_0}(\vec{r}_1)\,\mathrm{d}^3 r_1 \ , \tag{26.34}$$

et en itérant, on met en évidence un développement infini :

$$\psi_{\vec{k}_0}(\vec{r}) = \mathrm{e}^{\mathrm{i}k_0 z} + \sum_{n=1}^{+\infty}\left(\frac{-g}{4\pi}\right)^n \int_{\mathbb{R}^3}\ldots\int_{\mathbb{R}^3} G^{(+)}(\vec{r}-\vec{r}_1)G^{(+)}(\vec{r}_1-\vec{r}_2)\ldots G^{(+)}(\vec{r}_{n-1}-\vec{r}_n)\times \mathrm{e}^{\mathrm{i}k_0 z_n}\,\mathrm{d}^3 r_1\mathrm{d}^3 r_2\ldots\mathrm{d}^3 r_n \tag{26.35}$$

On ne sait pas en général en faire grand chose, mais on peut toujours définir des approximations successives en tronquant le développement, ce qui revient à écrire la solution à l'ordre n sous la forme :

$$\psi_{\vec{k}_0}^{(n)}(\vec{r}) = \mathrm{e}^{\mathrm{i}k_0 z} - g\,\frac{1}{4\pi}\int_{\mathbb{R}^3}\frac{\mathrm{e}^{\mathrm{i}k_0\|\vec{r}-\vec{r}\,'\|}}{\|\vec{r}-\vec{r}\,'\|}\,U(\vec{r}\,')\,\psi_{\vec{k}_0}^{(n-1)}(\vec{r}\,')\,\mathrm{d}^3 r' \ . \tag{26.36}$$

En règle générale, il n'est pas aisé de cerner les conditions de validité de ces développements tronqués, qui supposent d'ailleurs que la solution est analytique dans un voisinage de $g = 0$, autrement dit que la perturbation constituée par le potentiel n'est pas singulière. En tout cas, à l'ordre le plus bas, on a :

$$\psi_{\vec{k}_0}^{(1)}(\vec{r}) = \mathrm{e}^{\mathrm{i}k_0 z} - g\,\frac{1}{4\pi}\int_{\mathbb{R}^3}\frac{\mathrm{e}^{\mathrm{i}k_0\|\vec{r}-\vec{r}\,'\|}}{\|\vec{r}-\vec{r}\,'\|}\,U(\vec{r}\,')\,\mathrm{e}^{\mathrm{i}k_0 z'}\,\mathrm{d}^3 r' \ . \tag{26.37}$$

Cette approximation est dite *première approximation de Born* ; dans ce cadre, l'amplitude de diffusion est (voir (26.33)) :

$$\boxed{f_{k_0}^{(\mathrm{Born})}(\theta,\,\phi) = -g\,\frac{1}{4\pi}\int_{\mathbb{R}^3}\mathrm{e}^{-\mathrm{i}k_0\vec{n}.\vec{r}}\,U(\vec{r})\mathrm{e}^{\mathrm{i}k_0 z}\,\mathrm{d}^3 r} \tag{26.38}$$

Elle est calculable dès que l'on connaît la forme analytique du potentiel, et est essentiellement l'élément de matrice du potentiel entre les ondes (libres) incidente et diffusée. À un facteur près, c'est la transformée de Fourier du potentiel pour le vecteur d'onde $k_0.\vec{n} - (0, 0, k_0) \equiv \vec{k}_\mathrm{f} - \vec{k}_\mathrm{i}$, qui n'est autre que le *transfert de moment* lors de la collision[19] :

$$\boxed{f_{k_0}^{(\mathrm{Born})}(\theta,\,\phi) = -\frac{m}{2\pi\hbar^2}\mathcal{V}(\vec{k}_\mathrm{f}-\vec{k}_\mathrm{i}) \qquad \mathcal{V}(\vec{k}) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3}\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\,V(\vec{r})\,\mathrm{d}^3 r} \tag{26.39}$$

Ces expressions se simplifient quelque peu pour un potentiel central, puisqu'alors il est possible d'effectuer les intégrations sur les angles. Posant $\vec{q} \stackrel{\text{déf}}{=} \vec{k}_\mathrm{f} - \vec{k}_\mathrm{i}$, on obtient :

$$f_{k_0}^{(\mathrm{Born})}(\theta,\,\phi) = -\frac{2m}{\hbar^2 q}\int_0^{+\infty} rV(r)\sin qr\,\mathrm{d}r \ ; \tag{26.40}$$

[19] La collisison étant élastique, les deux vecteurs $\vec{k}_\mathrm{f}$ et $\vec{k}_\mathrm{i}$ ont le même module. Ils n'ont pas la même direction : comme discuté dans le Tome I, p. 61, une collision élastique sur une cible infiniment massive peut changer la direction de l'impulsion, mais pas son module.

comme les vecteurs $\vec{k}_\mathrm{f}$ et $\vec{k}_\mathrm{i}$ ont le même module (collision *élastique*) et font entre eux l'angle θ, on a $q = 2k_0 \sin\frac{\theta}{2}$, d'où l'amplitude de Born, *indépendante de* ϕ :

$$\boxed{f_{k_0}^{(\mathrm{Born})}(\theta) = -\frac{m}{\hbar^2 k_0 \sin\frac{\theta}{2}} \int_0^{+\infty} rV(r) \sin\big(2k_0 r \sin\frac{\theta}{2}\big)\,\mathrm{d}r} \tag{26.41}$$

À basse énergie, et si le potentiel est à courte portée, on a $k_0 a \ll 1$ pour toutes les valeurs importantes de l'intégrale ; remplaçant le sinus par son argument, l'amplitude prend la forme approximative :

$$f_{k_0}^{(\mathrm{Born})}(\theta) \simeq -\frac{2m}{\hbar^2} \int_0^{+\infty} r^2 V(r)\,\mathrm{d}r \ , \tag{26.42}$$

montrant que dans ces conditions, la section efficace est pratiquement indépendante de l'angle θ : à faible énergie, le projectile ne sait plus d'où il vient et sonde le potentiel à symétrie sphérique *assez longtemps* pour que toutes les directions se valent.

Au contraire, à haute énergie ($k_0 a \gg 1$), le sinus de l'intégrale oscille à toute vitesse, de sorte que celle-ci est pratiquement nulle sauf pour $\theta \simeq 0$ (voir problème 26.6.3) ; la présence du facteur $(\sin\frac{\theta}{2})^{-1}$ entraîne que, dans l'approximation de Born, l'amplitude diverge. Ceci est en tout point conforme à ce que l'on attend physiquement : à très haute énergie, le potentiel est quasi invisible et il n'y a pratiquement que de la diffusion vers l'avant. L'aspect directif est bien reproduit par le calcul approché ci-dessus ; quant à la divergence de l'amplitude au premier ordre, elle signifie simplement que celle-ci est infiniment grande par rapport à un résultat de type perturbatif... où par nature, la réponse au premier ordre est infiniment petite. On retiendra que, d'une façon générale et pour un potentiel à symétrie sphérique, la dépendance angulaire de la section efficace est d'autant plus forte que l'énergie des projectiles est élevée, un fait qui n'est pas surprenant à la réflexion.

26.4 Analyse en ondes partielles. Méthode des déphasages

Cette méthode est plus particulièrement conçue pour l'étude de la diffusion par un potentiel $V(r)$ à symétrie sphérique[20], et exploite la séparation des variables radiale et angulaires, suivant les résultats obtenus dans le chapitre 19. Dans ce chapitre, on a d'ailleurs déjà anticipé certains aspects de la diffusion, en examinant les états non-liés d'un potentiel central – voir par exemple le cas du puits "carré", sous-section 19.1.4. Quoi qu'il en soit, l'équation de base est l'équation aux fonctions propres qui, une fois effectuée la séparation $\psi(\vec{r}) = R(r)Y_{lm}(\theta, \phi)$ donne l'équation (19.38) pour la fonction

[20] Dès lors, l'amplitude f_k, et donc aussi la section efficace, ne dépend pas de l'angle azimutal ϕ, voir (26.41).

$u(r) \stackrel{\text{déf}}{=} rR(r)$, équation récrite ici sous la forme[21] :

$$\boxed{u''_{k_0l}(r) + \left[k_0^2 - U(r) - l(l+1)\frac{1}{r^2}\right] u_{k_0l}(r) = 0} \tag{26.43}$$

la fonction $u_{k_0l}(r)$ est paramétrée par l'entier $l \in \mathbb{N}$ et par le nombre d'onde $k_0 > 0$ donnant la valeur propre[22] du Hamiltonien sous la forme habituelle $E_0 = \frac{\hbar^2 k_0^2}{2m}$; $U(r)$ désigne la quantité[23] $\frac{2m}{\hbar^2}V(r)$.

On se place toujours dans le cas d'observations effectuées "à l'infini", ce qui signifie que pour un potentiel à support borné, les états propres sont ceux d'une particule libre dans $\mathbb{R}^3$, analysés en coordonnées sphériques (voir chapitre 19, sous-section 19.1.3, la particule libre est un cas particulier de champ central). Dans ce cas, les fonctions propres R_{k_0l} sont les fonctions de Bessel sphériques $j_l(k_0r)$, voir (19.84) ; supposant $V(r > a) \equiv 0$, la solution *exacte* satisfait (26.43) pour $r < a$, et, pour $r > a$:

$$u''_{k_0l}(r) + \left[k_0^2 - l(l+1)\frac{1}{r^2}\right] u_{k_0l}(r) = 0 \ , \tag{26.44}$$

Comme on l'a vu en détail dans l'exemple du puits carré (sous-section 19.1.4), la résolution effective contient notamment une étape de raccordement des deux familles de solutions à la surface du puits, laquelle constitue la frontière entre les régions intérieure et extérieure. Par rapport à la particule libre, les fonctions propres R_{k_0l} du puits carré pour $r > a$ diffèrent essentiellement par le fait qu'elles contiennent tout naturellement *deux* composantes, qui s'expriment à l'aide des fonctions de Bessel $J_{\pm(l+\frac{1}{2})}(k_0r)$ (voir (19.112)) – ces deux composantes traduisent précisément la possibilité de transmission et de réflexion par potentiel, pour employer le language (ici simpliste) des problèmes à une dimension. D'une façon générale, on a donc[24] :

$$r > a \ : \ \frac{1}{r}u_{k_0l}(r) \equiv R_{k_0l}(r) = A_{k_0l}j_l(k_0r) + B_{k_0l}y_l(k_0r) \ , \tag{26.45}$$

les coefficients A_{k_0l} et B_{k_0l} étant précisément fixés (à un facteur près) par les conditions de raccordement. Comme on l'a vu à propos du puits carré, leur variation en fonction de l'énergie est plutôt complexe (voir les figures 19.1 et 19.2).

Une fois les fonctions propres obtenues, leur développement asymptotique s'en déduit, connaissant[25] celui des fonctions de Bessel [38] :

$$k_0r \gg 1 \ : \ j_l(k_0r) \sim \frac{1}{k_0r}\sin\left(k_0r - l\frac{\pi}{2}\right) \ , \quad y_l(k_0r) \sim -\frac{1}{k_0r}\cos\left(k_0r - l\frac{\pi}{2}\right) \ ; \tag{26.46}$$

[21] Au prix d'une certaine lourdeur des écritures, mais par cohérence avec les notations précédentes, on continue à noter E_0 l'énergie initiale des particules, et k_0 le nombre d'onde correspondant.

[22] En quelque sorte, k_0 étiquette les états propres tout comme le fait un entier n pour les états liés du potentiel, quand ils existent, n donnant la valeur de l'énergie propre E_n de chaque état lié.

[23] Comme ce n'est pas utile ici, on ne fait pas apparaître une constante de couplage g, comme on l'avait fait antérieurement (comparer avec (26.5)).

[24] Les constantes A_{k_0l} et B_{k_0l} sont les analogues des coefficients $C_{E\pm}$ apparaissant dans l'expression (19.112), spécifique du puits carré.

Les fonctions j_l et y_l sont définies précisément dans le *Rappel sur les fonctions de Bessel* donné dans le problème 20.8.10 consacré au puits carré en théorie de Dirac.

[25] Ces expressions asymptotiques s'obtiennent en utilisant la méthode du col [13].

le report dans (26.45) donne une combinaison linéaire trigonométrique, que l'on peut toujours récrire avec une seule ligne trigonométrique déphasée, d'où la forme asymptotique (à un facteur près) de la fonction radiale :

$$\boxed{k_0 r \gg 1 \ : \quad \frac{1}{r} u_{k_0 l}(r) \equiv R_{k_0 l}(r) \sim \frac{1}{k_0 r} \sin\left[k_0 r - l\frac{\pi}{2} + \delta_l(k_0)\right]} \tag{26.47}$$

De toute évidence, le déphasage δ_l contient l'image du potentiel, ici carré, et dépend de l'énergie par l'intermédiaire de k_0.

Dans le cas d'un potentiel *quelconque*, à support non borné mais toujours caractérisable par une échelle de longueur encore notée a (portée finie), on ne connaît pas en général les fonctions propres exactes, mais il est plausible physiquement que leur comportement asymptotique ($r \gg a$) est de même nature[26] que ci-dessus si le potentiel décroît assez vite à l'infini, ce que l'on supposera. C'est pourquoi on peut espérer que, s'agissant d'obtenir les résultats dont on a pratiquement besoin en vue d'une expérience réelle de diffusion par un *bon* potentiel, le comportement asymptotique des solutions est toujours effectivement de la forme (26.47), où la quantité fondamentale à déterminer est le déphasage δ_l – d'où l'expression *méthode des déphasages* pour qualifier l'approche en cours. En repartant en temps utile de l'expression asymptotique déduite de l'équation intégrale de la diffusion, (26.32), on verra qu'il en est bien ainsi, encore que la démonstration manque quelque peu de rigueur mathématique ; en particulier, il conviendrait de valider soigneusement l'opération consistant, dans des *séries* de fonctions, à remplacer chacune de ces fonctions par son développement asymptotique pour obtenir le comportement asymptotique de la somme de la série. On admettra que ces substitutions sont toutes légitimes.

L'analyse en *ondes partielles* consiste à s'inspirer du développement d'une onde *plane* en ondes sphériques, donné en (19.91), récrit ici pour la commodité[27] :

$$\mathrm{e}^{\mathrm{i}k_0 r\cos\theta} = \sum_{l=0}^{+\infty} \mathrm{i}^l (2l+1)\, j_l(k_0 r) P_l(\cos\theta) \ , \tag{26.48}$$

où $P_l(\cos\theta)$ est le polynôme de Legendre de degré l (voir chapitre 18, section 18.3). En raison de (18.119) avec $m = 0$, on a :

$$P_l(\cos\theta) = \sqrt{\frac{4\pi}{2l+1}}\, Y_{l0}(\theta) \ , \tag{26.49}$$

d'où une autre écriture pour (26.48) :

$$\mathrm{e}^{\mathrm{i}k_0 z} = \sum_{l=0}^{+\infty} \mathrm{i}^l \sqrt{4\pi(2l+1)}\, j_l(k_0 r)\, Y_{l0}(\theta) \ . \tag{26.50}$$

[26]Cette affirmation est fausse pour le potentiel Coulombien, mais sans incidence sur la section efficace (voir Remarque p. 1213).
[27]Cette égalité est parfois appelée *formule de Bauer*.

Ces développements constituent la décomposition de $e^{ik_0 z}$ – qui, pour $k_0 > 0$, représente une onde plane (particule libre) avec un vecteur d'onde parallèle à Oz et dirigé dans le sens des z croissants –, fonction propre commune à $\vec{p}$ et $H_0 = \frac{\vec{p}^2}{2m}$, sur la base de fonctions propres de $(H_0, \vec{L}^2, L_z)$, lesquelles sont de la forme (voir (19.84)) :

$$\psi_{k_0 lm}^{(0)}(\vec{r}) \equiv R_{k_0 l}(r) Y_{lm}(\theta, \phi) = C_{k_0}\, j_l(k_0 r)\, Y_{lm}(\theta, \phi) \ , \tag{26.51}$$

où C_{k_0} est une constante de normalisation[28]. Les j_l dépendent de $E_0 = \frac{\hbar^2 k_0^2}{2m}$ par l'intermédiaire du nombre d'onde k_0, et forment une base complète. En utilisant le développement asymptotique des $j_l(k_0 r)$, (26.46), on peut écrire :

$$e^{ik_0 z} \sim \frac{1}{k_0 r} \sum_{l=0}^{+\infty} i^l (2l+1) \sin\left(k_0 r - l\frac{\pi}{2}\right) P_l(\cos\theta) \ , \tag{26.53}$$

En définitive, l'expression (26.53) est réputée[29] être la forme asymptotique (exacte) de la solution libre satisfaisant les conditions aux limites relatives à une source située à l'infini du côté $z < 0$. Noter que toute dépendance en ϕ disparaît puisque $V(r)$ est à symétrie sphérique, ce qui permet de remplacer les harmoniques sphériques Y_{l0} par les polynômes de Legendre, conformément à (26.49).

Introduisons maintenant le potentiel diffuseur ; dans le droit fil du développement de l'onde plane sur la base complète des états propres libres, on écrit de même pour la fonction propre $\psi_{\vec{k}_0}(\vec{r})$ d'énergie E_0 :

$$\psi_{\vec{k}_0}(\vec{r}) = \sum_{l=0}^{+\infty} a_l R_{k_0 l}(r) P_l(\cos\theta) \ , \tag{26.54}$$

où les coefficients a_l sont pour l'instant inconnus. Compte tenu de la forme symptotique (26.47) de chaque fonction radiale, l'expression pour $r \gg a$ de la fonction propre $\psi_{\vec{k}_0}(\vec{r})$ est :

$$\psi_{\vec{k}_0}(\vec{r}) \sim \sum_{l=0}^{+\infty} a_l \frac{1}{k_0 r} \sin\left(k_0 r - l\frac{\pi}{2} + \delta_l\right) P_l(\cos\theta) \ . \tag{26.55}$$

Revenons maintenant à l'équation (26.32), en y remplaçant le premier membre par son expression (26.55), et l'onde plane au second membre par la série (26.53) ; il vient de la sorte :

$$\sum_{l=0}^{+\infty} a_l \frac{1}{k_0 r} \sin\left(k_0 r - l\frac{\pi}{2} + \delta_l\right) P_l(\cos\theta) \sim \frac{1}{k_0 r} \sum_{l=0}^{+\infty} i^l (2l+1) \sin\left(k_0 r - l\frac{\pi}{2}\right) P_l(\cos\theta) + f_{k_0}(\theta) \frac{e^{ik_0 r}}{r} \ . \tag{26.56}$$

[28] Avec $C_k = \sqrt{2k^2/\pi}$, on a :

$$\langle \psi_{klm}^{(0)} | \psi_{k'l'm'}^{(0)} \rangle = \delta(k - k')\, \delta_{ll'}\, \delta_{mm'} \ . \tag{26.52}$$

[29] Comme dit plus haut, il conviendrait de justifier que le développement asymptotique de la somme de la série s'obtient bien en remplaçant chaque $j_l(k_0 r)$ par son développement asymptotique.

Pour fermer les équations, il convient maintenant d'exprimer l'amplitude de diffusion $f_{k_0}(\theta)$ à l'aide du développement asymptotique de $\psi_{\vec{k}_0}(\vec{r})$. D'après (26.33) (avec $g = 1$), on a ($\mathrm{d}^3r' = r'^2\mathrm{d}r'\sin\theta'\mathrm{d}\theta'\mathrm{d}\phi'$) :

$$f_{k_0}(\theta) = -\frac{1}{4\pi}\sum_{l=0}^{+\infty} a_l \int \mathrm{e}^{-\mathrm{i}k_0\vec{n}.\vec{r}'}U(r')P_l(\cos\theta')u_{k_0 l}(r')\, r'\sin\theta'\mathrm{d}\phi'\mathrm{d}r' \quad , \qquad (26.57)$$

où $\vec{n}$ est toujours le vecteur unitaire de l'argument $\vec{r}$ figurant dans $\psi_{\vec{k}_0}(\vec{r})$; désignant par Θ l'angle entre $\vec{r}$ et $\vec{r}'$, le facteur de phase sous l'intégrale est $\mathrm{e}^{-\mathrm{i}k_0 r'\cos\Theta}$. Afin d'effectuer commodément l'intégration sur les angles, on utilise à nouveau le développement (26.48) en écrivant :

$$\mathrm{e}^{-\mathrm{i}k_0 r'\cos\Theta} = \sum_{l'=0}^{+\infty}\mathrm{i}^{-l'}(2l'+1)\, j_{l'}(k_0 r')P_{l'}(\cos\Theta) \ ; \qquad (26.58)$$

injectant ceci dans l'intégrale, on obtient notamment l'intégrale angulaire :

$$\int_0^{2\pi}\mathrm{d}\phi' \int_0^{\pi}\mathrm{d}\theta'\, P_{l'}(\cos\Theta)P_l(\cos\theta')\sin\theta' \quad , \qquad (26.59)$$

qui se calcule facilement en utilisant d'abord le théorème d'addition des harmoniques sphériques, égalité (18.121), puis leur relation d'orthonormalisation, (18.73). On trouve ainsi que l'intégrale (26.59) est égale à $\frac{4\pi}{2l+1}P_l(\cos\theta)\,\delta_{ll'}$, d'où, selon (26.57) :

$$f_{k_0}(\theta) = -\sum_{l=0}^{+\infty}\mathrm{i}^{-l}a_l P_l(\cos\theta)\int_0^{+\infty} j_l(k_0 r')U(r')u_{k_0 l}(r')\, r'\mathrm{d}r' \quad , \qquad (26.60)$$

Il se trouve que l'intégrale radiale s'exprime simplement en fonction du déphasage δ_l. En effet, considérons d'abord la fonction $u_{k_0 l}^{(0)}(r) \stackrel{\text{déf}}{=} rj_l(k_0 r)$, qui est la solution correspondant à la particule libre et satisfait donc (26.43) avec $U \equiv 0$. Désignant toujours par $u_{k_0 l}$ la solution propre pour U non nul, multiplions membre à membre (26.43) par $u_{k_0 l}^{(0)}$, cette même équation à U nul par $u_{k_0 l}$ et retranchons-les ; de la sorte, on trouve :

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[u_{k_0 l}^{(0)}u'_{k_0 l} - u'^{(0)}_{k_0 l}u_{k_0 l}\right] = u_{k_0 l}^{(0)}U u_{k_0 l} \quad ; \qquad (26.61)$$

en intégrant de 0 à $+\infty$, la valeur en $r = 0$ disparaît en vertu des propriétés de la fonction radiale de type $u(r)$ (voir (19.60)). La valeur limite supérieure en $+\infty$ s'obtient par manipulation des formes asymptotiques (26.47), et on trouve qu'elle vaut $-\frac{1}{k_0}\sin\delta_l$, d'où l'égalité importante :

$$\boxed{\sin\delta_l(k_0) = -k_0\int_0^{+\infty} j_l(k_0 r)U(r)u_{k_0 l}(r)\, r\mathrm{d}r} \qquad (26.62)$$

cette intégrale est très précisément celle qui figure dans (26.60), d'où :

$$f_{k_0}(\theta) = \frac{1}{k_0}\sum_{l=0}^{+\infty}\mathrm{i}^{-l}a_l\sin\delta_l\, P_l(\cos\theta) \ . \qquad (26.63)$$

Reportant cette expression au second membre de (26.56) et utilisant les formules d'Euler, on obtient des développements sur les fonctions $r^{-1}\mathrm{e}^{\mathrm{i}kr}P_l(\cos\theta)$; en identifant les coefficients d'un membre à l'autre, on trouve finalement les a_l sous la forme :

$$a_l = \mathrm{i}^l(2l+1)\,\mathrm{e}^{\mathrm{i}\delta_l} \ , \tag{26.64}$$

achevant la détermination de la forme asymptotique (26.54) de la solution propre d'énergie E_0 :

$$\boxed{\psi_{\vec{k}_0}(\vec{r}\,) \sim \frac{1}{k_0 r}\sum_{l=0}^{+\infty} \mathrm{i}^l(2l+1)\,\mathrm{e}^{\mathrm{i}\delta_l}\sin(k_0 r - l\frac{\pi}{2} + \delta_l)\,P_l(\cos\theta)} \tag{26.65}$$

On en déduit également le résultat central de la méhode des déphasages, donnant l'amplitude de diffusion en fonction de ces derniers :

$$\boxed{f_{k_0}(\theta) = \frac{1}{k_0}\sum_{l=0}^{+\infty}(2l+1)\,\mathrm{e}^{\mathrm{i}\delta_l}\sin\delta_l\,P_l(\cos\theta)} \tag{26.66}$$

Compte tenu de (26.49), on a aussi :

$$f_{k_0}(\theta) = \frac{1}{k_0}\sum_{l=0}^{+\infty} \sqrt{4\pi(2l+1)}\;\mathrm{e}^{\mathrm{i}\delta_l}\;\sin\delta_l\,Y_{l0}(\theta) \ . \tag{26.67}$$

Ces résultats sont tout à fait remarquables, au sens où ils sont valides pour tout *bon* potentiel, quelle qu'en soit la forme précise, et où ils donnent l'expression asymptotique *exacte* de l'onde diffusée. Remarquable aussi est la *forme* du développement (26.65) : par comparaison avec le développement de l'onde plane (26.53) (solution à $V(r) \equiv 0$), on voit que la présence du potentiel conduit à des changements minimalistes (en apparence) : la seule modification visible est l'apparition du facteur de phase $\mathrm{e}^{\mathrm{i}\delta_l}$ signifiant que chaque composante sphérique l conserve son nombre quantique l et subit seulement un glissement de phase mesuré par δ_l. Physiquement, le potentiel déforme l'onde incidente pour lui imprimer son image : cette distortion se traduit par le fait que chaque composante sphérique acquiert son propre déphasage, dépendant aussi tout naturellement de l'énergie E_0 des projectiles.

D'après (26.16) et par intégration sur les angles selon (26.3), on en déduit la section efficace totale σ ; compte tenu de l'orthonormalisation des $Y_{l\,m}$, il vient :

$$\boxed{\sigma(k_0) = \frac{4\pi}{k_0^2}\sum_{l=0}^{+\infty}(2l+1)\sin^2\delta_l(k_0) \stackrel{\text{déf}}{=} \sum_{l=0}^{+\infty}\sigma_l} \tag{26.68}$$

En pratique, un tel développement n'est vraiment utile que pour les potentiels donnant un petit nombre de déphasages sensiblement différents de zéro[30]. Les δ_l peuvent en principe se trouver par identification, si le potentiel $V(r)$ est connu, en résolvant

[30] La question n'est pas tant de sommer la série que de savoir calculer de nombreux déphasages.

l'équation aux valeurs propres correspondante pour chaque valeur de l et en écrivant l'expression asymptotique des fonctions propres. Noter que les amplitudes des différentes ondes partielles apparaissant dans (26.67) s'additionnent de façon *incohérente* pour donner la section efficace totale, au sens où, en vertu de l'orthogonalité des harmoniques sphériques, $\sigma(k_0)$ s'obtient en faisant la somme des modules au carré des termes de la série donnant l'amplitude de diffusion, le facteur de phase $\mathrm{e}^{\mathrm{i}\delta_l}$ disparaissant *ipso facto.*

Il se trouve que le nombre de déphasages pertinents peut être semi-quantitativement relié à la portée spatiale du potentiel. En effet, l'expression intégrale des déphasages, (26.62), permet de les évaluer au premier ordre non nul en remplaçant $u_{k_0 l}(r)$ par sa valeur à l'ordre zéro, soit $rj_l(k_0 r)$:

$$\sin\delta_l \simeq -k\int_0^{+\infty}\left[j_l(k_0 r)\right]^2 U(r)\, r^2 \mathrm{d}r \ . \tag{26.69}$$

L'idée est d'estimer cette intégrale en profitant du fait que les fonctions $j_{l>0}$ partent toutes de zéro et *décollent* très lentement quand $l \gg 1$; plus précisément, on a :

$$j_l(k_0 r) = \frac{(k_0 r)^l}{(2l+1)!!} + \ldots \ , \qquad (k_0 r \ll 1) \tag{26.70}$$

qui montre que plus l est grand, plus la fonction j_l est "écrasée" dans le voisinage de l'origine ; un calcul précis[31] montre que j_l décolle pour $k_0 r \sim l$, ce qui permet de définir une longueur typique $\sim l/k$ au-dessous de laquelle la fonction $j_l(k_0 r)$ est pratiquement nulle. Inversement, ceci signifie que pour une énergie donnée (k_0 fixé), les déphasages de grand l seront tous presque nuls puisque là où la fonction j_l est sensiblement non-nulle, le potentiel est nul. Plus précisément, si on définit l'entier $l_{\mathrm{max}} \stackrel{\text{déf}}{=} \mathrm{E}[k_0 a]$, les déphasages $\delta_{l>l_{\mathrm{max}}}$ seront tous très petits. Le nombre de déphasages utiles est donc d'autant plus petit que l'énergie est basse :

$$\boxed{l \lesssim l_{\mathrm{max}} \sim k_0 a = \hbar^{-1}\sqrt{2mE_0}\, a = 2\pi\frac{a}{\lambda_{\mathrm{dB}}}} \tag{26.71}$$

et, sans surprise, il augmente avec l'énergie : à haute énergie, la diffusion est très fortement axiale dans la direction incidente et un grand nombre de termes dans les développements en ondes sphériques est requis. À l'inverse, plus précisément pour un potentiel de portée a petite par rapport à λ_{dB} (diffusion à basse énergie, $E \propto (\lambda_{\mathrm{dB}})^{-/2}$), *un seul* sera déphasage sensiblement non-nul, δ_0.

Ainsi, souvent, l'amplitude (26.66) est donnée à une bonne approximation par une somme contenant un très petit nombre de termes. L'habitude a été donnée d'affubler du symbole S, P, D, F,... les termes correspondants à l = 0, 1, 2, 3, ... ; on écrit ainsi (sachant que $P_1(\cos\theta) = \cos\theta$) :

$$f_{k_0}^{\mathrm{S}} = \frac{1}{k_0}\mathrm{e}^{\mathrm{i}\delta_0}\sin\delta_0 \ , \quad f_{k_0}^{\mathrm{P}} = \frac{3}{k_0}\mathrm{e}^{\mathrm{i}\delta_1}\sin\delta_0\cos\theta \ , \quad \text{etc.} \tag{26.72}$$

[31] Utiliser la formule de Stirling $n! \simeq \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$, qui donne déjà des résultats remarquables pour les petites valeurs de n ($5! = 120$, $(5!)_{\mathrm{Stirling}} \simeq 118,019\ldots$).

et aussi :

$$f_{k_0}^{\mathrm{S+P}} = \frac{1}{k_0}\left(\mathrm{e}^{\mathrm{i}\delta_0}\,\sin\delta_0 + 3\mathrm{e}^{\mathrm{i}\delta_1}\,\sin\delta_0\cos\theta\right)\ . \tag{26.73}$$

Les sections efficaces correspondantes sont spécifiées de la même façon.

Un résultat important mérite d'être mentionné : le signe de tous les déphasages est déterminé par la nature attractive ou répulsive du potentiel ; ils sont tous positifs dans le premier cas, négatifs dans le second, ce qui se comprend assez bien physiquement. Pour un potentiel attractif, le projectile a tendance à *traîner* dans le voisinage du potentiel, et prend donc du retard ; au contraire, il prend plutôt de l'avance quand le potentiel est répulsif.

Pour démontrer ce résultat, récrivons l'équation aux fonctions propres (26.43) en réintroduisant la constante de couplage sans dimension $g > 0$ en facteur du potentiel :

$$u_{k_0 l}''(r) + \left[k_0^2 - gU(r) - l(l+1)\frac{1}{r^2}\right] u_{k_0 l}(r) = 0 \tag{26.74}$$

Écrivons maintenant cette équation pour deux valeurs de g, g_1 et g_2, et soit $u_{k_0 l\, g_i}$ ($i = 1,\, 2$) les deux fonctions correspondantes. Maintenant, multiplions l'équation avec g_1 membre à membre par $u_{k_0 l\, g_2}$, celle avec g_2 par $u_{k_0 l\, g_1}$, et retranchons-les. En suivant exactement la même procédure que celle conduisant à (26.62), on obtient :

$$\frac{1}{g_2 - g_1}\left[\sin\delta_l(g_2) - \sin\delta_l(g_1)\right] = -k\int_0^{+\infty} u_{k_0 l\, g_1} U(r) u_{k_0 l\, g_2}\,\mathrm{d}r\ ; \tag{26.75}$$

prenant maintenant la limite $g_2 \to g_1$, on obtient :

$$\boxed{\frac{\partial}{\partial g}\,\sin\delta_l(g) = -k_0\int_0^{+\infty} U(r)[u_{k_0 l\, g}]^2\,\mathrm{d}r} \tag{26.76}$$

Dans la limite $g \to 0_+$, $u_{k_0 l\, g} \to j_l(k_0 r)$, et l'égalité (26.62) montre alors que $\sin\delta_l$ est positif ou nul[32] si $U < 0$ quel que soit r, et inversement. Cela étant, avec $U < 0$, (26.76) montre que la dérivée de δ_l est positive : dans ces conditions, δ_l part d'une valeur positive et croît quand la force du potentiel augmente. Le contraire se produit si le potentiel est répulsif quel que soit r. Au total, on a bien le résultat annoncé :

$$\boxed{\forall l\ :\quad V(r)\ \text{répulsif} \iff \delta_l \leq 0\ ,\quad V(r)\ \text{attractif} \iff \delta_l \geq 0} \tag{26.77}$$

À titre d'illustration, considérons un exemple simple, qui conduit d'ailleurs à un résultat un peu surprenant quand on le confronte avec le cas classique : la diffusion par un anti-puits infini sphérique (sphère dure). Dans ce cas :

$$V(r) = \begin{cases} 0 & \text{si } r > a \\ +\infty & \text{si } r < a \end{cases}\ . \tag{26.78}$$

[32]L'équation (26.62) définit le déphasage par son sinus ; on peut, conventionnellement et sans affecter la signification physique, décréter que δ_l tend vers zéro quand $g \to 0$; c'est par valeurs positives si $U < 0$, négatives si $U > 0$.

D'après ce qui précède, seuls les déphasages satisfaisant (26.71) sont pertinents. En choisissant une énergie suffisamment faible (donc k_0 petit), on peut se placer dans les conditions où seule l'onde S ($l = 0$) compte. Il reste alors, suivant (26.67) :

$$f_{k_0}(\theta) \simeq \mathrm{e}^{\mathrm{i}\delta_0(k_0)} \sin\delta_0(k_0) \ . \tag{26.79}$$

f_{k_0} est indépendant de l'angle de diffusion (une conséquence du fait que seule l'onde sphérique importe vraiment) ; la section efficace totale est :

$$\sigma \simeq \frac{4\pi}{k_0^2} \sin^2\delta_0(k_0) \ . \tag{26.80}$$

Il reste à trouver $\delta_0(k_0)$, ce que l'on peut faire en résolvant l'équation radiale avec $l = 0$. La fonction $u_{k_0 0}(r) = r\,R_{k_0 0}(r)$ satisfait :

$$u''_{k_0 0} + k_0^2\, u_{k_0 0} = 0 \qquad (\forall r > a) \ , \tag{26.81}$$

avec $u_{k_0 0}(r = a) = 0$ (le puits est impénétrable). La solution est :

$$u_{k_0 0}(r) = \begin{cases} A\,\sin[k_0(r-a)] & \text{si } r > a \\ 0 & \text{si } r < a \end{cases} \ . \tag{26.82}$$

Par ailleurs, le déphasage est identifié en écrivant le comportement asymptotique de la fonction radiale $R_{k_0 0}(r)$ (voir (26.55) et (26.64)) :

$$R_{k_0 0}(r) \simeq \frac{1}{k_0 r}\,\mathrm{e}^{\mathrm{i}\delta_0} \sin(k_0 r + \delta_0) \iff u_{k_0 0}(r) \simeq C\,\sin(k_0 r + \delta_0) \quad (r \rightarrow +\infty) \ . \tag{26.83}$$

Par comparaison avec (26.82), $\delta_0 = -k_0 a$ et donc, suivant (26.68), $\sigma = \frac{4\pi}{k_0^2}\sin^2 k_0 a$. En fait, comme on a supposé $k_0 a \ll 1$, le résultat significatif est en réalité :

$$\boxed{\sigma \simeq 4\pi a^2} \tag{26.84}$$

Ce résultat, valide notamment à très basse énergie, montre que dans une telle limite la section efficace est indépendante de l'énergie. L'expression (26.84) n'est en fait rien d'autre que la section efficace totale classique ($\hbar = 0$) d'une sphère dure de rayon a pour un projectile ponctuel... à un facteur 4 près ! ($\sigma_{\hbar=0} = \pi\, a^2$). On peut voir dans cette différence un effet quantique irréductible lié à la variation brutale du potentiel sur une échelle de longueur toujours plus petite que toute autre longueur[33], puisque V a un saut (variation sur une longueur nulle). En réalité, il s'agit encore d'un cas où la limite classique doit être analysée avec soin[34].

D'une façon générale, la section efficace est d'autant plus petite que l'énergie est élevée et que l'entier l est grand. Il existe toutefois, tout comme le cas unidimensionnel

[33] comme par exemple la longueur d'onde de de Broglie $h/\sqrt{2mE_0} = 2\pi/k_0$.

[34] On se trouve dans la situation analogue à celle rencontrée à propos d'une marche de potentiel, voir Tome I p. 567 : l désignant la (petite) longueur sur laquelle le potentiel passe de 0 à l'infini, il faut en réalité examiner la commutation des deux limites $\hbar \rightarrow 0$ et $l \rightarrow 0$. Une autre "anomalie" du même type a aussi été rencontrée pour les deux limites $\hbar \rightarrow 0$ et $T \rightarrow 0$ (voir Tome I, note 42 p. 622).

(voir Tome I, p. 556), des résonances de diffusion à basse énergie, liées au fait que, la particule se déplaçant lentement, elle a bien le temps de sonder le potentiel. Ces résonances, visibles dès l'onde P ($l = 1$), sont examinées en détail dans la section suivante.

Auparavant, il convient de noter une relation très simple entre la section efficace totale et la partie imaginaire de l'amplitude de diffusion. Comme $P_l(0) = 1$ quel que soit l'entier l, l'amplitude en $\theta = 0$ vaut $f_{k_0}(0) = \frac{1}{k_0}\sum_l (2l+1)\,\mathrm{e}^{\mathrm{i}\delta_l}\sin\delta_l$; prenant la partie imaginaire, il vient :

$$\Im f_{k_0}(0) = \frac{1}{k_0}\sum_{l=0}^{+\infty}(2l+1)\,\sin^2\delta_l \quad ; \tag{26.85}$$

la comparaison avec la section efficace totale (26.68) montre que :

$$\boxed{\sigma = \frac{4\pi}{k_0}\Im f_{k_0}(0)} \tag{26.86}$$

Cette relation s'appelle *théorème optique* (*optical theorem*), ainsi nommé par analogie avec la relation, en Optique, entre la partie imaginaire de l'indice de réfraction et le coefficient d'absorption [195].

26.5 Résonances de diffusion

Il s'agit ici de préciser les conditions donnant lieu à une variation rapide, avec l'énergie, de certains déphasages autour de la valeur $\frac{\pi}{2}$, produisant des *résonances* de la section efficace au voisinage de certaines énergies. Afin de simplifier la discussion, on suppose que le potentiel est à support borné, non-nul pour $r < a$, strictement nul pour $r > a$. Dans ces conditions, la solution *exacte*, pour $r > a$, a l'expression asymptotique donnée par (26.45) soit, une fois introduit le déphasage δ_l (voir (26.47)) :

$$R_{k_0 l} = \cos\delta_l\, j_l(k_0 r) - \sin\delta_l\, y_l(k_0 r) \qquad (r > a)\ . \tag{26.87}$$

Il reste, pour un potentiel de forme prélablement précisée, à résoudre l'équation aux fonctions propres pour $r < a$ puis à écrire la continuité de la fonction radiale et de sa dérivée en $r = a$. Cette condition de raccordement s'exprime simplement à l'aide du rapport $R'_{k_0 l}/R_{k_0 l}$ (dérivée logarithmique). Soit $R^{\mathrm{int}}_{k_0 l}$ la solution à l'intérieur du puits d'énergie E_0 et de moment angulaire l ; introduisons la quantité sans dimension $\beta_l(k_0)$

$$\beta_l(k_0) \stackrel{\text{déf}}{=} a\left(\frac{1}{R^{\mathrm{int}}_{k_0 l}}\frac{\mathrm{d}R^{\mathrm{int}}_{k_0 l}}{\mathrm{d}r}\right)_{r=a} \quad ; \tag{26.88}$$

on peut montrer (voir problème 26.6.4) que $\beta_l(k_0)$ est une fonction *décroissante* de k_0, donc de l'énergie E_0 du projectile. Cela étant, la condition de raccordement à la frontière du puits s'écrit :

$$k_0 a\,\frac{\cos\delta_l\, j'_l(k_0 a) - \sin\delta_l\, y'_l(k_0 a)}{\cos\delta_l\, j_l(k_0 a) - \sin\delta_l\, y_l(k_0 a)} = \beta_{k_0 l}\ , \tag{26.89}$$

et se résout facilement pour $\tan\delta_l$:

$$\boxed{\tan\delta_l(k_0) = \frac{k_0 a\, j_l'(k_0 a) - \beta_l(k_0)\, j_l(k_0 a)}{k_0 a\, y_l'(k_0 a) - \beta_l(k_0)\, y_l(k_0 a)}} \qquad (26.90)$$

La possibilité de résonances de diffusion réside dans le fait que le dénominateur peut devenir très petit sur un intervalle d'énergie très étroit, grâce à un ajustement très fin de la dérivée logarithmique $\beta_l(k_0)$ par rapport au produit $k_0 a y_l'/y_l$; le plus souvent, et comme les fonctions de Bessel ont une variation relativement lente, le cœur du phénomène se tient dans une variation brutale de la dérivée logarithmique $\beta_l(k_0)$.

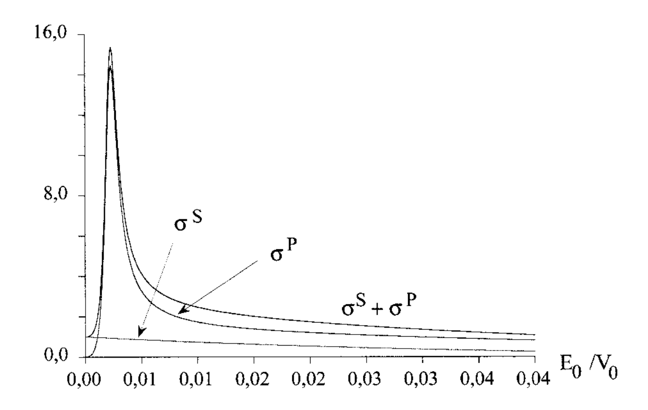


Figure 26.3: Variation en fonction de l'énergie incidente E_0 des sections efficaces σ^{S}, σ^{P} et de leur somme pour un puits "carré" de rayon a et de profondeur V_0 modérée (les courbes sont tracées avec $a\hbar^{-1}\sqrt{2mV_0} \simeq 10$). L'unité en ordonnée est $4\pi a^2$, limite de σ^{S} quand $V_0 \to \pm\infty$. La résonance basse énergie de σ^{P} est très visible.

Dans le cas très particulier mais néanmoins exemplaire du puits carré, les calculs explicites peuvent être conduits jusqu'au bout ; il suffit ici d'exploiter les résultats obtenus dans la sous-section 19.1.4 à propos de l'étude détaillée des états non-liés d'un tel puits. La connaissance des relations précises entre les coefficients alors notés[35] $C_{E\pm}$ donne immédiatement le déphasage δ_l (voir éqs. (19.116) et (19.117)), et permet de tracer exactement les sections efficaces σ^{S} et σ^{P} (les autres étant très petites). La figure 26.3 permet de mesurer la brusque variation de σ^{P}, autorisant à parler réellement de *résonance*.

On peut donner une description semi-quantitative des résonances en abandonnant le cas particulier du puits carré. Afin de simplifier les écritures, on introduit les fonctions

[35] Noter également que la quantité notée alors k_0 est liée à la profondeur du puits par $k_0 \stackrel{\text{déf}}{=} \hbar^{-1}\sqrt{2mV_0}$, et non à l'énergie incidente comme dans la section en cours.

dites de Hankel[36] :

$$h_l^{(\pm)}(z) \stackrel{\text{déf}}{=} j_l(z) \pm \mathrm{i} y_l(z) \ ; \tag{26.91}$$

cela étant, il se révèle commode de définir un certain déphasage ξ_l par :

$$\mathrm{e}^{2\mathrm{i}\xi_l} \stackrel{\text{déf}}{=} -\frac{h_l^{(-)}(k_0 a)}{h_l^{(+)}(k_0 a)} \ . \tag{26.92}$$

Avec ces définitions, un calcul élémentaire montre que la condition de raccordement (26.90) prend la forme :

$$\mathrm{e}^{2\mathrm{i}\delta_l} = \left(1 + k_0 a \frac{B_l^{(+)} - B_l^{(-)}}{\beta_l - k_0 a B_l^{(+)}}\right) \mathrm{e}^{2\mathrm{i}\xi_l} \ , \tag{26.93}$$

où les $B_l^{(\pm)}$ sont les dérivées logarithmiques des fonctions $h_l^{(\pm)}$, évaluées pour $r = a$. Cette égalité permet déjà de donner un sens physique aux déphasages ξ_l. En effet, si $\beta_l = \infty$, (26.93) montre que les déphasages δ_l et ξ_l sont égaux ; or β_l diverge quel que soit l pour un potentiel de sphère dure puisque, en pareil cas, la fonction propre est strictement nulle sur la frontière du puits. Autrement dit, les ξ_l ne sont rien d'autre que les déphasages pour une sphère dure de rayon a. Posant enfin :

$$ka\, B_l^{(+)}(k_0 a) \stackrel{\text{déf}}{=} \Delta_l + \mathrm{i}\gamma_l \ , \tag{26.94}$$

l'égalité (26.93) s'écrit comme suit :

$$\boxed{\mathrm{e}^{2\mathrm{i}(\delta_l - \xi_l)} = \frac{\beta_l - \Delta_l + \mathrm{i}\gamma_l}{\beta_l - \Delta_l - \mathrm{i}\gamma_l}} \tag{26.95}$$

La quantité γ_l est positive, en conséquence du fait que le Wronskien[37] des fonctions de Bessel sphériques, $j_l(z)y_l'(z) - j_l'(z)y_l(z)$, est égal à z^{-2} ([38], § 8.474). Pour le puits carré, les différentes quantités introduites ci-dessus ont une expression simple dans le cas de l'onde S ($l = 0$) ; on trouve $\xi_0 = -k_0 a$, $\Delta_0 = -1$ et $\gamma_0 = ka$; en outre, avec $K_0 \stackrel{\text{déf}}{=} \hbar^{-1}\sqrt{2mV_0}$, $K \stackrel{\text{déf}}{=} \sqrt{k_0^2 + K_0^2}$, on a $\beta_0 = Ka \cot Ka - 1$, qui est bien une fonction décroissante de l'énergie.

Afin d'appréhender qualitativement l'origine de la résonance, admettons que la dérivée logarithmique – dont c'est la variation rapide qui engendre la résonance, le plus souvent – peut être raisonnablement remplacée par sa variation linéaire dans le très petit intervalle d'énergie centré sur la résonance. Dans ces conditions, et choisissant de faire figurer l'énergie incidente E_0 plutôt que le nombre d'onde k_0, on écrit $\beta_l \simeq b - b'E_0$, avec $b' > 0$ puisque β_l est une fonction décroissante de l'énergie ; l'égalité (26.95) prend alors la forme :

$$\mathrm{e}^{2\mathrm{i}(\delta_l - \xi_l)} \simeq \frac{E_0 - E_{l\mathrm{r}} - \frac{\mathrm{i}}{2}\hbar\Gamma_l}{E_0 - E_{l\mathrm{r}} + \frac{\mathrm{i}}{2}\hbar\Gamma_l} \ , \tag{26.96}$$

[36] Ces fonctions sont parfois notées $h_l^{(i)}(z)$ ($i = 1, 2$). Ici, tous les arguments étant réels, on a simplement $h_l^{(2)}(x) = [h_l^{(1)}(x)]^*$.

[37] Cette propriété a déjà été utilisée à propos du puits carré (sous-section 19.1.4).

où $E_{l\text{r}}$ et $\Gamma > 0$ s'expriment simplement en fonction de b, b', Δ_l et γ_l. On en déduit

$$\tan(\delta_l - \xi_l) \simeq \frac{\hbar\Gamma}{2(E_{l\text{r}} - E_0)} \ . \tag{26.97}$$

Clairement, cette approximation est cohérente si l'assimilation linéaire pour $\beta_l(k_0)$ est bonne sur un intervalle autour de $E_{l\text{r}}$ assez grand devant la largeur $\hbar\Gamma$. Admettons enfin que le déphasage ξ_l varie lentement près de $E_{l\text{r}}$; l'expression (26.97) montre que, dans ces conditions, δ_l varie rapidement entre les deux valeurs $\simeq \xi_l \pm \frac{\pi}{2}$ quand E_0 franchit la valeur $E_{l\text{r}}$; la variation rapide du déphasage produit une variation brutale de la section efficace σ_l correspondante. Le report dans l'expression de σ_l (voir (26.68)) donne :

$$\boxed{\sigma_l(E_0) \simeq (2l+1)\frac{4\pi}{k_0^2}\frac{(\hbar\Gamma/2)^2}{(E_0 - E_{l\text{r}})^2 + (\hbar\Gamma/2)^2}} \tag{26.98}$$

expression typique d'une résonance, d'autant plus fine que la largeur Γ est petite. Tout comme dans le cas unidimensionnel, les résonances sont la traduction, du côté des énergies positives, des états liés du potentiel. Il faut cependant garder en tête les approximations faites pour arriver à l'expression (26.98), et en particulier l'hypothèse d'une variation lente pour le déphasage de sphère dure ξ_l, aussi nanti d'une valeur *banale*. En fait, si ξ_l est alors voisin de $(2n+1)\frac{\pi}{2}$, $n \in \mathbb{Z}$, ... la section efficace tombe à zéro à la valeur qui, sans cette coïncidence, donnerait une résonance. On retiendra que la détection des résonances, tout comme leur interprétation, sont parfois des opérations assez subtiles.

Par ailleurs, il a été mentionné que la Mécanique quantique autorise l'existence d'états liés (normalisables) noyés dans le continuum des états de diffusion [35] (voir p. 784). En termes de résonance de diffusion, cette situation assez exotique se traduit par des résonances de diffusion *infiniment* fines.

♦ *Remarque*

Tout au long de ce chapitre, et de façon récurrente, on s'est appuyé sur le fait que le potentiel diffuseur était de portée finie et bien identifiable, ce qui excluait *ipso facto* le cas du potentiel de Coulomb. Ce cas est important, même s'il est finalement assez rare en pratique en raison des inévitables phénomènes d'écrantage. Quoi qu'il en soit, pour ce potentiel qui reste remarquable, il est possible d'effectuer les calculs jusqu'au bout, au prix d'une certaine complexité technique ([10], § 133, [53], V.20). Tout comme pour l'effet Stark, l'existence d'une direction privilégiée (la direction du faisceau incident) suggère l'utilisation des coordonnées paraboliques, ce qui permet de résoudre exactement l'équation aux fonctions propres[38] (qui n'est pas paramétrée par l'entier l puisque l'on n'est pas en coordonnées sphériques), et d'obtenir une expression relativement simple des fonctions propres (on rencontre à nouveau la fonction hypergéométrique). Une fois cette solution exacte obtenue, le retour en sphériques permet d'obtenir, par un développement asymptotique, l'onde

[38]Il est aussi possible de rester en coordonnées sphériques d'un bout à l'autre ; de la sorte, on obtient la fonction radiale $u_{k_0}(r)$ sous la forme (pas très agréable !) d'une *série* de fonctions hypergéométriques ([53], éq. (20.21)).

diffusée à grande distance dans le système de coordonnées approprié à l'analyse de l'onde sphérique.

Le changement majeur, par rapport à la situation standard d'un potentiel à courte portée, est la dépendance *radiale* de l'onde diffusée. Notant $V(r) = \frac{qQ}{4\pi\varepsilon_0 r}$ le potentiel et $\gamma \stackrel{\text{déf}}{=} \frac{mqQ}{\hbar^2 k_0}$, le terme contenant l'amplitude de diffusion $f_{k_0}(\theta)$ s'écrit :

$$\frac{1}{r} f_{k_0}(\theta)\, \mathrm{e}^{\mathrm{i}[k_0 r - \gamma \ln(2k_0 r)]} \equiv \frac{1}{(2\mathrm{i}k_0)^{\mathrm{i}\gamma}} \frac{1}{r^{1+\mathrm{i}\gamma}} f_{k_0}(\theta)\, \mathrm{e}^{\mathrm{i}k_0 r} \ . \tag{26.99}$$

La distorsion logarithmique se retrouve également dans l'onde diffusée vers l'avant : ainsi, en raison de la portée infinie du potentiel de Coulomb nu, les dépendances radiales sont affectées aussi bien pour l'onde transmise que pour l'onde sphérique. Si ce changement qualitatif est remarquable, il n'apporte que des *corrections* au courant, lesquelles tendent heureusement[39] vers zéro à l'infini et sont donc finalement inessentielles car sans effet sur le taux de comptage à très grande distance, donc sur la section efficace [209]. En définitive, et malgré la portée infinie, la section efficace est encore donnée par l'expression (26.16). ♦

26.6 Exercices et problèmes

26.6.1 Équation intégrale de la diffusion dans $\mathbb{R}$

La diffusion par un potentiel à une dimension a été rencontrée dans le Tome I (voir chapitre 15), et traitée par les moyens élémentaires. Il s'agit ici de reformuler cette question en utilisant, cette fois dans $\mathbb{R}$, l'équation intégrale de la diffusion présentée pour $\mathbb{R}^3$ dans la section 26.3, dont on reprend les notations.

1. Récrire l'équation dans $\mathbb{R}$ pour les états non-liés du potentiel $V(x)$ (équivalent unidimensionnel de (26.18)). En déduire l'équation intégrale.

2. La source étant placée en $x = -\infty$ comment se comporte la solution $\psi_{k_0}(x)$ en $x = \pm\infty$?

3. En suivant la même procédure que dans la section 26.3, montrer que la bonne fonction de Green est ici :

$$G(x,\, x') = \frac{1}{2\mathrm{i}k_0}\, \mathrm{e}^{\mathrm{i}k_0|x-x'|} \ . \tag{26.100}$$

4. En déduire l'expression formelle de la solution ψ_{k_0} sous une forme analogue à (26.20).

[39]Voir les commentaires dans la note 9 p. 1194.

5. Appliquer ces résultats au cas du potentiel de Dirac, $V(r) = -g\delta(x)$, et comparer aux résultats du problème 15.6.5 du Tome I.

6. Revenant à un potentiel quelconque, écrire l'expression intégrale formelle des coefficients de transmission, T, et de réflexion, R.

7. Examiner la première approximation de Born.

8. Développer le même formalisme au cas des états *liés*, en posant $E = -\frac{\hbar^2 k^2}{2m}$, $k > 0$. En particulier, montrer que la bonne fonction de Green est :

$$G(x, x') = -\frac{1}{2k}\,\mathrm{e}^{-k|x-x'|} \ , \tag{26.101}$$

et en déduire l'équation intégrale satisfaite par une fonction propre $\psi_k(x)$.

9. Appliquer ce qui précède au cas de $V(x) = -g\delta(x)$ et montrer qu'il n'existe de solution non-identiquement nulle que si $k = \frac{mg}{\hbar^2}$. Commenter en comparant aux résultats du problème 15.6.5 du Tome I.

26.6.2 Absence de diffusion pour le potentiel de Dirac

Montrer que tous les déphasages δ_l sont nuls pour le potentiel sphérique $V(r) = g\delta(r)$.

26.6.3 Analyse de l'approximation de Born en fonction de l'énergie

Il s'agit de discuter plus en détail l'expression (26.41) donnant l'amplitude de diffusion élastique $f_{k_0}^{(\mathrm{Born})}(\theta)$ dans la (première) approximation de Born, pour des particules de masse m et d'énergie initiale $E_0 = \frac{\hbar^2 k_0^2}{2m}$. On se place dans le cas ordinaire d'un potentiel central à portée a bien identifiée ; afin de préciser les choses, on pourra faire des calculs explicites pour les deux potentiels modèles[40] :

$$V_{\mathrm{Y}}(r) \stackrel{\text{déf}}{=} V_0 \frac{a}{r}\,\mathrm{e}^{-\mu r/a} \ , \quad V_{\mathrm{G}}(r) = V_0 \mathrm{e}^{-(r/a)^2} \qquad (\mu > 0, a > 0,\, V_0 \in \mathbb{R}) \ ; \tag{26.102}$$

Dans tous les cas, on introduira la fonction $v(x = \frac{r}{a}) \stackrel{\text{déf}}{=} V(r)$.

1. Montrer que, à très basse énergie, la section efficace est pratiquement indépendante de l'angle de diffusion.

[40] V_{Y} est le potentiel de Yukawa.

2. On se place maintenant à haute énergie, en posant $\lambda \stackrel{\text{déf}}{=} 2k_0 a$. Argumenter pour montrer que l'amplitude peut s'écrire sous la forme :

$$f_{k_0}^{(\text{Born})}(\theta) = -\frac{2ma^3}{\hbar^2\lambda^3 \sin\frac{\theta}{2}} \int_0^{+\infty} [v(0) + \frac{X}{\lambda}v'(0) + ...] \sin(X\sin\frac{\theta}{2})\,\mathrm{d}X \ . \quad (26.103)$$

3. Montrer que l'intégrale de (26.103) se comporte essentiellement comme $2\pi\delta(\theta)$.

4. Commenter ce résultat, et argumenter physiquement pour comprendre le fait que, indépendamment de l'"infini" introduit par la fonction de Dirac, on trouve une divergence de la section efficace.

5. Calculer effectivement $f_{k_0}^{(\text{Born})}(\theta)$ pour les potentiels gaussien et de Yukawa. Dans ce dernier cas, examiner la limite $\mu \to 0$.

6. Commenter le signe de $f_{k_0}^{(\text{Born})}(\theta)$ en fonction du caractère attractif ou répulsif du potentiel.

26.6.4 Décroissance avec l'énergie de la dérivée logarithmique β_l définie en (26.88)

On reprend les notations de la section 26.5 ; en particulier, le potentiel $V(r)$ est supposé non nul pour $0 < r < a$.

1. Soit deux solutions stationnaires $\Psi_i(\vec{r}, t) \stackrel{\text{déf}}{=} \mathrm{e}^{\frac{1}{i\hbar}E_i t}\psi_i(\vec{r})$ de l'équation de Schrödinger ; montrer que :

$$\frac{\partial}{\partial t}\big(\Psi_1^*(\vec{r},t)\Psi_2(\vec{r},t)\big) + \frac{\hbar}{2im}\vec{\nabla}.\big[\Psi_1^*(\vec{r},t)\vec{\nabla}\Psi_2(\vec{r},t) - \big(\vec{\nabla}\Psi_1^*(\vec{r},t)\big)\Psi_2(\vec{r},t)\big] = 0 \ . \quad (26.104)$$

2. Appliquer cette égalité au cas où $\psi_i(\vec{r}) = R_{E_i l}(r)Y_{lm_i}(\theta, \phi)$ pour en déduire :

$$(E_2 - E_1)r^2 R_{E_1 l}(r)R_{E_2 l}(r) + \frac{\hbar^2}{2m}\frac{\mathrm{d}}{\mathrm{d}r}\Big\{r^2\big[R_{E_1 l}(r)R'_{E_2 l}(r) - R'_{E_1 l}(r)R_{E_2 l}(r)\big]\Big\} = 0 \ . \quad (26.105)$$

3. Par une intégration de cette équation entre 0 et a, montrer que :

$$\beta_l(E_2) - \beta_l(E_1) = -\frac{2m(E_2 - E_1)}{\hbar^2 a R_{E_1 l}(a)R_{E_2 l}(a)} \int_0^a R_{E_1 l}(r)R_{E_2 l}(r)r^2\,\mathrm{d}r \qquad (26.106)$$

4. En déduire que $\beta_l(E)$ est une fonction monotone décroissante de l'énergie E.

26.6.5 Diffusion de l'onde S par un puits carré

Il s'agit d'illustrer en détails certains résultats de ce chapitre pour le puits carré sphérique $V(r) = -V_0$ si $r \leq a$, $V(r) = 0$ si $r > a$. La fonction de Bessel sphérique $l = 0$ est $j_0(\rho) = \frac{\sin\rho}{\rho}$. On pose $K \stackrel{\text{déf}}{=} \sqrt{\frac{2m}{\hbar^2}(E_0 + V_0)} \equiv \sqrt{k_0^2 + K_0^2}$, où $E_0 = \frac{\hbar^2 k_0^2}{2m}$ est l'énergie des projectiles.

1. Écrire l'expression de la dérivée β_0.
2. Quelles sont des fonctions de Hankel $h_0^{(\pm)}(\rho)$?
3. Trouver le déphasage ξ_0 défini en (26.92).
4. Combien vaut la dérivée logarithmique $B_0^{(+)}$ apparaissant dans (26.93) ? Trouver les paramètres Δ_0 et γ_0 définis par (26.94).
5. En déduire l'amplitude de diffusion $f^{\text{S}}(\theta)$, puis la section efficace $\sigma^{\text{S}}(\theta)$.
6. Montrer que, à basse énergie :

$$\sigma^{\text{S}}(\theta) \simeq \Big(\frac{\tan K_0 a}{K_0 a} - 1\Big)^2 . \tag{26.107}$$

   À l'opposé, comment varie $\sigma^{\text{S}}(\theta)$ à haute énergie ?
7. On définit[41] la *longueur de diffusion* l_{diff} comme :

$$l_{\text{diff}} \stackrel{\text{déf}}{=} -\lim_{k\to 0} \frac{\delta_0}{k} . \tag{26.108}$$

   Étudier son signe, et faire le rapprochement avec la condition d'existence d'états liés du puits[42].
8. Que se passe-t-il pour $K_0 = (2n+1)\frac{\pi}{2}$?
9. Les énergies des résonances sont $E_r = -V_0 + (2n+1)^2 \frac{\pi^2\hbar^2}{8ma^2}$. Expliciter l'approximation linéaire $\beta_0 \simeq b - b' E_0$.
10. En déduire la variation résonnante de la section efficace $\sigma_0(E)$ pour $E_0 \sim E_{\text{r}}$ (voir (26.98)).

26.6.6 Approximation de Born pour le puits carré

Soit le potentiel $V(r) = -V_0$ si $r < a$, $V(r) = 0$ si $r > a$. Trouver la section efficace différentielle $\sigma_{\text{d}}(\theta)$ dans la première approximation de Born et tracer sa variation en fonction de θ.

[41] Suivant le contexte et les auteurs, l_{diff} est parfois définie avec le signe contraire.

[42] Ce résultat est général : pour un potentiel à symétrie sphérique, la longueur de diffusion ainsi définie est négative s'il n'existe aucun état lié de type S ($l = 0$) ([207], p. 178).

26.6.7 Miracle de l'approximation de Born pour le potentiel Coulombien nu

En utilisant certains résultats du problème 26.6.3, montrer que l'approximation de Born restitue l'expression exacte de la section efficace pour la diffusion Rutherford (voir Tome I, éq. (3.51))

26.6.8 Approximation de Born pour le potentiel de Yukawa

Dans l'approximation de Born, trouver les sections efficaces différentielle et totale pour le potentiel $V(r) = -\frac{g}{r}\mathrm{e}^{-\lambda r}$. La section efficace totale est-elle finie ? Retrouver le potentiel Coulombien à la limite.

26.6.9 Densité en champ moyen pour un gaz d'électrons

Soit un gaz de N électrons confinés dans une boîte cubique de coté L. On admet que, dans une étape préalable, on a traité l'interaction entre électrons par une procédure de type champ moyen, permettant de définir l'énergie potentielle moyenne d'un électron dans le champ des autres, soit $V(\vec{r})$. Dans la suite, on décrit la dynamique du gaz en traitant, dans l'approximation de Born, la diffusion d'un électron sur le champ effectif créé par les autres.

1. Montrer que la fonction d'onde[43] $\psi_{\vec{k}}(\vec{r})$ d'un électron d'énergie $E = \frac{\hbar^2 k^2}{2m}$ satisfait l'équation intégrale :

$$\psi_{\vec{k}}(\vec{r}) = L^{-3/2}\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} - \frac{m}{2\pi\hbar^2}\int_{\mathbb{R}^3} \frac{1}{\|\vec{r}-\vec{r}\,'\|}\,\mathrm{e}^{\mathrm{i}k\|\vec{r}-\vec{r}\,'\|}\,V(\vec{r}\,')\psi_{\vec{k}}(\vec{r}\,')\,\mathrm{d}^3r' \ . \qquad (26.109)$$

2. Expliquer pourquoi la densité électronique est donnée par l'expression :

$$\rho(\vec{r}) = 2\sum_{k\leq k_\mathrm{F}} \psi^*_{\vec{k}}(\vec{r})\psi_{\vec{k}}(\vec{r}) \qquad (26.110)$$

Préciser qualitativement le sens du nombre d'onde noté k_F.

3. Dans la suite, se restreint à la première approximation de Born. Écrire l'équation intégrale satisfaite par $\psi_{\vec{k}}(\vec{r})$.

4. Montrer que, à cet ordre, la densité est donnée par :

$$\rho(\vec{r}) = \rho_0 - L^{-3}\frac{m}{\pi\hbar^2}\Big[\sum_{k\leq k_\mathrm{F}} \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\,I(\vec{k}) + \mathrm{c.c.}\Big] \ , \qquad (26.111)$$

où $I(\vec{k})$ est une certaine intégrale à préciser, et $\rho_0 \stackrel{\text{déf}}{=} \frac{N}{L^3}$.

[43] Il s'agit en fait d'une *orbitale*.

5. Effectuer la substitution $L^{-3}\sum_{k\leq k_\mathrm{F}} \to \frac{1}{(2\pi)^3}\int_{\|\vec{k}\|\leq k_\mathrm{F}} \mathrm{d}^3k$ pour obtenir $\rho(\vec{r})$ sous la forme :

$$\rho(\vec{r}) = \rho_0 - \frac{m}{\pi\hbar^2}\int \frac{1}{\|\vec{r}-\vec{r}\,'\|} V(\vec{r}\,')\Omega(\|\vec{r}-\vec{r}\,'\|)\,\mathrm{d}^3r' \ . \tag{26.112}$$

6. Calculer explicitement l'intégrale définissant la fonction $\Omega(\|\vec{r}-\vec{r}\,'\|)$, et montrer qu'elle s'exprime à l'aide de la fonction de Bessel sphérique :

$$j_1(x) \stackrel{\text{déf}}{=} \frac{1}{x^2}(\sin x - x\cos x) \ . \tag{26.113}$$

7. On pose $E_0 \stackrel{\text{déf}}{=} \frac{\hbar^2}{2m}\rho_0^{2/3}$ et on admet[44] que k_F est proportionnel à $\rho_0^{1/3}$. En déduire que la densité $\rho(\vec{r})$ est donnée par :

$$\rho(\vec{r}) = \rho_0\left[1 - \frac{C}{E_0}\rho_0^{1/3}\int \frac{j_1(2k_\mathrm{F}\|\vec{r}-\vec{r}\,'\|)}{\|\vec{r}-\vec{r}\,'\|^2} V(\vec{r}\,')\,\mathrm{d}^3r'\right] \ , \tag{26.114}$$

où C est un certain nombre.

8. Montrer que loin du point $\vec{r}$ à l'échelle $(2k_\mathrm{F})^{-1}$, l'effet du potentiel est atténué par un facteur $\sim \frac{\cos 2k_\mathrm{F}\|\vec{r}-\vec{r}\,'\|)}{\|\vec{r}-\vec{r}\,'\|^3}$ (oscillations de Friedel).

[44] Pouvez-vous en trouver une brève justification ?

Partie IV

Applications à basse énergie

physique

Chapitre 27

Atomes à plusieurs électrons

Le traitement des atomes complexes est présenté en partant du modèle à électrons indépendants et en précisant les notions de terme spectral et de multiplet, uniquement liées à la symétrie de rotation. Les structures fine et hyperfine sont ensuite abordées. La dernière section est consacrée au cas le plus simple (l'hélium), pour lequel les calculs sont conduits en détail jusqu'au bout afin de mettre en lumière le rôle crucial de la répulsion de Coulomb pour l'apparition éventuelle d'un moment magnétique atomique.

Pour un atome de charge nucléaire Z et possédant N électrons, le Hamiltonien électrostatique est, avec des notations évidentes (et en supposant le noyau infiniment massif) :

$$H_{\mathrm{e}} = \sum_{i=1}^{N} \left(\frac{\vec{p}_i^{\,2}}{2m} - \frac{Ze'^2}{r_i}\right) + \frac{1}{2}\sum_{i=1}^{N}\ \sum_{j=1,\, j\neq i}^{N} \frac{e'^2}{r_{ij}} \equiv \sum_{i=1}^{N} H_1(i) + V \ . \qquad (27.1)$$

Le dernier terme représente la répulsion Coulombienne des électrons entre eux et constitue un terme essentiellement déstabilisant (le facteur $\frac{1}{2}$ est là pour ne pas compter deux fois l'interaction au sein d'une même paire), dont on prévoit qu'il est de l'ordre de quelques eV et est donc loin d'être petit. D'ailleurs, même s'il l'était, il faut réaliser qu'il y a $N(N-1) \sim N^2$ termes de cette sorte ; ceux-ci, tous positifs, sont donc beaucoup plus nombreux que les $N \sim Z$ termes représentant l'attration du noyau sur les électrons. Dans tous les cas, le terme à deux corps joue donc visiblement un rôle important. Par ailleurs, les termes de structure fine (voir section 27.3), généralisant à N électrons ce qui a été introduit dans le chapitre 20, contiennent *a priori* des termes à deux corps, du genre $\vec{l}_i.\vec{s}_j$ et $\vec{s}_i.\vec{s}_j$, qui sont toutefois beaucoup plus petits que V. En règle générale, le Hamiltonien

d'un atome complexe est donc de la forme :

$$H(\vec{x}_1, \vec{x}_2, \dots \vec{x}_N) = \sum_{i=1}^{N} H_1(\vec{x}_i) + \frac{1}{2} \sum_{i,j=1}^{N} H_2(\vec{x}_i, \vec{x}_j) \ , \tag{27.2}$$

où $\vec{x}_i$ désigne l'ensemble des variables pour un électron (coordonnées et spin), et où H_1 et H_2 sont des opérateurs à un et deux corps respectivement.

À cause des termes à deux corps, on ne sait pas résoudre exactement l'équation aux valeurs propres et il faut recourir à des descriptions approchées – que l'on ne peut d'ailleurs mettre réellement en œuvre dans les cas intéressants qu'en utilisant l'outil numérique. L'atome à N électrons est un premier exemple de système relevant des techniques du problème à N corps (N- *body*, ou *many-body, problem*), omniprésent en Physique et de toute première importance puisque c'est quand on *branche* les interactions que des propriétés remarquables peuvent surgir. Ainsi par exemple, c'est l'interaction *électrostatique* entre deux électrons au sein d'un atome qui détermine les propriétés *magnétiques* de ce dernier. En outre, les interactions peuvent engendrer des comportements *collectifs* pour une assemblée de particules – par *collectif*, on entend une modification *qualitative* par rapport aux comportements prévisibles en simple réunion des comportements individuels. Un exemple spectaculaire de ce type de transition est la théorie BCS de la supraconductivité : si les deux électrons candidats à une paire de Cooper n'étaient pas *noyés* dans le fluide des autres électrons, cette paire ne pourrait se former (voir chapitre 32, section 32.5).

Les petits systèmes (atomes à quelques électrons) ont été l'un des premiers terrains d'expérimentation de la Mécanique quantique, et ont permis de dessiner des méthodes d'approximation qui ont fait école dans presque tous les autres problèmes où les interactions entre particules (fermions notamment) jouent un rôle décisif. Par exemple, le modèle de Heisenberg (voir chapitre 30, sous-section 30.6.1, éq. 30.202) pour des spins couplés est construit sur la généralisation à N spins de l'interaction effective que l'on peut construire pour un atome à deux électrons (voir problème 27.8.1).

27.1 Modèle à électrons indépendants

Ce modèle est défini en adoptant l'approximation la plus triviale, consistant à oublier purement et simplement la répulsion V dans (27.1), et tous les autres termes à deux corps éventuels. Ce procédé brutal conduira à des résultats très médiocres, quantitativement bien sûr, mais aussi qualitativement – ce qui permettra d'ailleurs de montrer à quel point l'interaction (alliée au Principe de Pauli) joue un rôle de tout premier plan. Malgré tout, le modèle n'est pas sans intérêt dans la mesure où il permet d'introduire des idées généralisables dans des approches plus subtiles et aussi parce que, finalement, il constitue le fondement le plus simple du principe de la classification périodique, au moins dans ses grandes lignes. Enfin, sa discussion permet de mettre en lumière le fait fondamental que deux particules (identiques) indépendantes au sens classique sont malgré tout corrélées

en raison de leur indiscernabilité ; une illustration fameuse de cette *interaction quantique* est le trou de Fermi déjà rencontré au chapitre 22 (voir éq. (22.85) et problème 22.6.1).

Dans ces conditions, le Hamiltonien à considérer est :

$$H_{\text{ind}}(\vec{x}_1, \vec{x}_2, \dots \vec{x}_N) = \sum_{i=1}^{N} H_1(\vec{x}_i) \ , \tag{27.3}$$

H_{ind} est une somme de Hamiltoniens à une particule, $H_1(\vec{x}_i)$, tous identiques (les électrons sont identiques !) et ne différant dans la sommation que par l'indice i des variables dont ils dépendent et sur lesquelles ils agissent. Désignons maintenant par ψ_k et E_k les modes propres de $H_1(\vec{x})$:

$$H_1(\vec{x})\psi_k(\vec{x}) = E_k\psi_k(\vec{x}) \ . \tag{27.4}$$

Si H_1 ne dépend pas du spin, $H_1(\vec{x}) \equiv H_1(\vec{r}, \vec{p})$, la fonction ψ_k peut toujours être prise comme le produit d'une fonction ϕ_k par une fonction de spin χ, ϕ_k étant propre de H_1 ; en représentation-q :

$$\psi_k(\vec{x}) \equiv \phi_k(\vec{r})\chi(m_s) \ , \qquad H_1(\vec{r}, -\mathrm{i}\hbar\vec{\nabla})\phi_k(\vec{r}) = E_k\phi_k(\vec{r}) \ . \tag{27.5}$$

Suivant la terminologie déjà introduite au chapitre 22, la fonction $\psi_k(\vec{x})$ est une spin-orbitale, la fonction $\phi_k(\vec{r})$ est une orbitale ; dans la suite, on continue à raisonner en toute généralité avec les spin-orbitales, dans la mesure du possible.

H_{ind} est une somme d'opérateurs H_1 à un corps : comme on l'a vu précédemment (chapitre 22, section 22.4), tout produit de spin-orbitales ψ_k propres de H_1 est fonction propre de H_{ind} avec pour valeur propre la somme des valeurs propres associées aux spin-orbitales :

$$H \prod_{j=1}^{N} \psi_{k_j}(\vec{x}_j) = \Big(\sum_{i=1}^{N} E_{k_i}\Big) \prod_{j=1}^{N} \psi_{k_j}(\vec{x}_j) \tag{27.6}$$

Cela étant, le produit $\prod_{j=1}^{N} \psi_{k_j}$ n'est pas une fonction physiquement acceptable vis-à-vis du postulat de symétrisation : les électrons sont des fermions et la transposition des variables de deux quelconques d'entre eux doit donner un changement de signe dans la fonction d'onde. Suivant la prescription énoncée dans le chapitre 22, il faut appliquer au produit de spin-orbitales l'opérateur d'antisymétrisation Y_{A} défini en (22.43), qui est un projecteur, afin d'obtenir une fonction ayant la bonne symétrie de permutation :

$$P_\mu Y_{\text{A}}\Phi = (-1)^{\sigma_\mu}\, Y_{\text{A}}\Phi \ , \tag{27.7}$$

Si Φ est le produit de spin-orbitales ci-dessus, la fonction $\Psi \stackrel{\text{déf}}{=} Y_{\text{A}}\Phi$ est une fonction acceptable : elle est fonction propre de H_{ind} – permuter les variables n'altère pas la qualité de fonction propre de H_{ind} – et satisfait les exigences liées à l'indiscernabilité des particules identiques. En définitive, toute fonction propre du modèle à électrons indépendants est de la forme :

$$\boxed{\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = C\, Y_{\text{A}} \prod_{j=1}^{N} \psi_{k_j}(\vec{x}_j)} \tag{27.8}$$

où C est une constante de normalisation. La normalisation s'écrit :

$$1 = \langle\Psi|\Psi\rangle = |C|^2\langle Y_\mathrm{A}\Phi|Y_\mathrm{A}\Phi\rangle = |C|^2\langle\Phi|Y_\mathrm{A}\Phi\rangle \ , \tag{27.9}$$

où la troisième égalité vient de l'hermiticité et de l'idempotence de Y_A. Si les spin-orbitales sont orthonormalisées, toute permutation dans le produit scalaire à droite donne zéro, à l'exception de la permutation identité, d'où $C = \sqrt{N!}$. Pour un atome de numéro atomique Z, la logique du modèle à électrons indépendants suggère de prendre comme orbitales les fonctions propres hydrogénoïdes ; c'est visiblement le choix le plus simpliste dans un modèle qui l'est déjà.

Le modèle à électrons indépendants conduit naturellement au modèle en couches de l'atome. Chaque niveau hydrogénoïde E_n est dégénéré[1] n^2 fois ; compte tenu du spin, on peut donc mettre $2n^2$ électrons[2] dans l'ensemble $\{\psi_{nlmm_s}\}_{lmm_s}$. La couche K ($n = 1$) peut donc contenir 2 électrons, et contient une seule sous-couche, notée 1s ($l = 0$). La couche L ($n = 2$) contient les deux sous-couches 2s et 2p ($l = 1$) qui peuvent accommoder respectivement 2 et 6 électrons. La couche M regroupe les trois sous-couches s, p et d ($l = 2$) et peut accueillir $2 + 6 + 10 = 18$ électrons, et ainsi de suite. Au moins au début de la classification, l'oubli de la répulsion électrostatique n'est pas catastrophique et l'ordonnancement des énergies hydrogénoïdes suffit à comprendre qualitativement la constitution de l'état fondamental des atomes pas trop lourds, au moins en terme de configuration électronique. Ainsi, l'état fondamental du carbone (6 électrons), est caractérisé par la configuration électronique[3] $(1\mathrm{s})^2(2\mathrm{s})^2(2\mathrm{p})^2$. D'un autre côté, en l'absence de toute interaction – s'en tenant donc strictement aux niveaux hydrogénoïdes – on ne comprend pas pourquoi un atome aussi simple que le lithium ($Z = 3$) a pour configuration fondamentale $(1\mathrm{s})^2(2\mathrm{s})$ et pourquoi pas $(1\mathrm{s})^2(2\mathrm{p})$: ces deux configurations ont, de ce point de vue, strictement la même énergie ; le même argument vaut pour le bore ($Z = 4$), dont la configuration fondamentale est $(1\mathrm{s})^2(2\mathrm{s})^2$, laquelle a la même énergie que $(1\mathrm{s})^2(2s)(2p)$, $(1\mathrm{s})^2(2p)^2$, etc.

27.2 Termes spectraux

Que l'on prenne ou non en compte la répulsion entre les électrons, il est possible de faire une présentation générale reposant exclusivement sur la symétrie du Hamiltonien, restant pour l'instant dans le strict cadre électrostatique. Comme le système est *globalement* invariant par rotation, le moment cinétique total[4] est une constante du mouvement.

[1] Se souvenir de la dégénérescence "accidentelle" du champ Coulombien nu.

[2] L'énergie d'un état hydrogénoïde étant $E_n = -\frac{Z^2\alpha^2}{2n^2}mc^2$, l'énergie de liaison d'une couche complète, hors répulsion des électrons entre eux, est exactement égale à $Z^2\alpha^2mc^2$. Elle est en fait beaucoup plus petite, compte tenu justement des interactions entre électrons, qui est déstabilisante.

[3] Cette configuration ne permet pas de comprendre la valence 4 du carbone ; on rend compte de celle-ci en invoquant la notion d'hybridation, consistant à prendre une orbitale 2s et à la combiner avec les 3 orbitales p (hybridation notée $(\mathrm{sp})^3$ par les chimistes) ; la valence 3 (comme dans le benzène, voir chapitre 28, sous-section 28.2.4) s'explique par une hybridation sp^2.

[4] H ne commute évidemment pas avec la somme partielle des moments cinétiques d'une partie des électrons : il n'y a invariance par rotation que si *tous* les électrons tournent en bloc.

De surcroît, dans l'approximation électrostatique (délaissant donc la structure fine), H commute évidemment avec tous les opérateurs de spin et en particulier avec le spin total $\vec{S} = \sum_{i=1}^{N} \vec{s}_i$. En définitive, il existe des vecteurs propres communs à H, $\vec{L}^2$, L_z, $\vec{S}^2$ et S_z, et ce, quel que soit le traitement que l'on envisage pour traiter au mieux l'interaction entre les électrons[5].

Dans l'hypothèse où la notion de configuration électronique garde toute sa pertinence, le moment cinétique orbital total s'obtient en combinant les moments cinétiques $\vec{l}_i$ des orbitales, représentés par les nombres quantiques l_i et m_{l_i}, $1 \leq i \leq N$. La discussion générale est sans grand intérêt, mais on peut noter qu'il est facile de savoir quelles valeurs de L sont associées à une configuration donnée $\ldots\, (n_i,\, l_i)^{p_i}\, \ldots$, sachant que $M_L = \sum_{i=1}^{N} m_{l_i}$. Notons d'abord que la somme (partielle) des m_{l_i} relatifs à une couche complète est nulle, donc on ne peut trouver d'états $L \neq 0$ – qui exigeraient $M_L \neq 0$; il suffit donc de regarder les couches incomplètement remplies et de trouver les valeurs possibles de la somme des m_{l_i} associés, sans oublier le principe de Pauli. Ayant trouvé le M_L maximum, $M_{L,\,\mathrm{max}}$, on peut affirmer que les valeurs de L – moment cinétique *orbital* – sont $M_{L,\,\mathrm{max}}$, $M_{L,\,\mathrm{max}} - 1$, ... 0. Il n'y a évidemment qu'une seule façon de combiner les m_{l_i} pour obtenir $M_{L,\,\mathrm{max}}$ (chacun d'entre eux prend sa valeur maximale l_i) ; en revanche, il y a plusieurs façons de construire la somme égale à $M_{L,\,\mathrm{max}} - 1$. L'une de ces combinaisons donnera la composante $M_L - 1$ associée à $L_{\mathrm{max}} \equiv M_{L,\,\mathrm{max}}$, les autres – orthogonales – donneront les composantes maximum pour $L_{\mathrm{max}} - 1$. On retrouve ici, mais appliquée systématiquement de proche en proche, la méthode mise en œuvre pour l'addition de *deux* moments cinétiques (voir chapitre 18, sous-section 18.5.2).

La même procédure vaut pour le spin total $\vec{S} = \sum_{i=1}^{N} \vec{s}_i$, puisque chaque s_i prend exclusivement la valeur $\frac{1}{2}$. Ainsi, pour N électrons, la valeur maximale de S est :

$$\boxed{S_{\mathrm{max}} = \frac{N}{2}} \tag{27.10}$$

La composante $M_S = S$ (état parfois dit (abusivement !) *ferromagnétique*) associée à S_{max} s'obtient en accolant chaque orbitale une fois occupée avec la fonction de spin α. La fonction d'onde correspondante se forme ainsi à partir de[6] :

$$\Phi_{S_{\mathrm{max}}\, M_S = S_{\mathrm{max}}} = \left[\prod_{i=1}^{N} \phi_{n_i l_i m_{l_i}}(\vec{r}_i)\right] \prod_{j=1}^{N} \alpha(m_{s_j}) \ . \tag{27.11}$$

La bonne fonction antisymétrique correspondante est alors :

$$\Psi_{S_{\mathrm{max}}\, M_S = S_{\mathrm{max}}} = Y_{\mathrm{A}}\left[\Phi_{S_{\mathrm{max}}\, M_S = S_{\mathrm{max}}}\right] \ ; \tag{27.12}$$

comme la partie de spin de l'expression (27.11) est totalement symétrique, il vient simplement :

$$\Psi_{S_{\mathrm{max}}\, M_S = S_{\mathrm{max}}}(\vec{x}_1,\, \vec{x}_2,\, \ldots\, \vec{x}_N) = \left[Y_{\mathrm{A}} \prod_{i=1}^{N} \phi_{n_i l_i m_{l_i}}(\vec{r}_i)\right] \prod_{j=1}^{N} \alpha(m_{s_j}) \ . \tag{27.13}$$

[5] Contrairement à ce que l'on voit parfois écrit, il n'y a pas lieu d'émettre une quelconque hypothèse de *champ central*.

[6] Les deux fonctions $\alpha(m)$ et $\beta(m)$ ont été définies dans le chapitre 20, éq. (20.36).

Pour l'état de plus haute multiplicité de spin, la fonction d'onde une fois antisymétrisée est encore[7] le produit d'une fonction d'espace par une fonction de spin[8], une factorisation qui reste visiblement vraie pour les $2S_{\text{max}}+1$ composantes (S_- n'agit que sur le facteur de spin), mais qui ne se produit plus dès que $S < S_{\text{max}}$ – pour ces états, les variables d'espace et de spin sont forcément imbriquées (*intriquées*, au sens de ce mot défini dans le chapitre 21, section 21.5). Maintenant, en faisant agir l'opérateur de descente $S_- = \sum_{i=1}^{N} s_{i\,-}$, on obtient la composante $M_{S_{\text{max}}} - 1$ de l'état $S_{\text{max}} = \frac{N}{2}$. Compte tenu de la forme de la fonction de spin dans (27.13) et de :

$$s_+|\alpha\rangle = s_-|\beta\rangle = 0 \ , \qquad s_-|\alpha\rangle \propto |\beta\rangle \ , \qquad s_+|\beta\rangle \propto |\alpha\rangle \ , \tag{27.14}$$

cette composante apparaît sous la forme d'une combinaison linéaire de N termes, contenant chacun $(N-1)$ fonctions α et une fonction β (il y a N façons de basculer un spin) ; on peut donc affirmer qu'il existe $N-1$ états de spin $S = S_{\text{max}} - 1 \geq 0$. Plus généralement, il n'est pas difficile de montrer qu'il y a $\text{C}_N^p - \text{C}_N^{p-1}$ états de spin $S = S_{\text{max}} - p$ (voir problème 18.7.8) ; la somme de toutes les multiplicités de spin doit bien évidemment être égale à 2^N, ce que l'on vérifie trivialement en additionnant les différences $\text{C}_N^p - \text{C}_N^{p-1}$, dont la somme reconstitue $(1+1)^N$ par le développement du binôme.

En définitive, choisir une valeur de L et une valeur de S revient à considérer au total $(2L+1)(2S+1)$ états propres distincts constituant ce que l'on appelle un *terme spectral,* noté comme suit :

$$^{2S+1}\mathcal{F}(L) \qquad (\mathcal{F}(L) = \text{S, P, D, F,} \ldots \text{ si } L = 0, 1, 2, 3, \ldots) \ . \tag{27.15}$$

Ainsi, un terme spectral $S = 1$, $L = 2$ est noté ^{3}D ; de même, l'état fondamental d'un atome ayant toutes ses couches complètes et un électron célibataire dans une orbitale s (l'argent par exemple) sera du type ^{2}S. Pour distinguer deux termes de même symétrie, l'usage est d'ajouter un numéro d'ordre ; ainsi, on note 1^2S l'état fondamental de l'hydrogène et 2^2S le premier état excité ayant $L = 0$. Enfin, rappelons que dans l'approximation à électrons indépendants, tous les termes spectraux issus d'une même configuration électronique construite sur des orbitales hydrogénoïdes sont dégénérés.

27.3 Structure fine. Multiplets

Tout ce qui a été dit précédemment concerne exclusivement l'approximation électrostatique. Une description plus précise doit prendre en compte l'interaction spin-orbite responsable, en tant que petite correction[9] (pour les éléments légers), de la structure fine (éclatement de chaque terme spectral en plusieurs composantes formant un multiplet).

La bonne façon d'obtenir l'ensemble des termes décrivant la structure fine consiste à partir de l'équation de Dirac pour les N électrons (4^N composantes !) et à en prendre

[7] dans l'approximation électrostatique.

[8] Cette propriété est toujours vraie dans le cas de $N = 2$ électrons.

[9] Le rapport entre l'interaction spin-orbite et la différence d'énergie entre deux termes spectraux est typiquement de l'ordre de α^2, les corrections relativistes augmentant, en gros, comme Z^2.

la limite faiblement relativiste pour obtenir le Hamiltonien de Pauli. Ce travail est fait en détail dans le livre de Bethe et Salpeter ([89], § 39), et conduit à une expression assez compliquée (voir l'équation (39.14) de cet ouvrage), dont les termes les plus importants (pour la structure fine) sont brièvement décrits ci-dessous.

Le Hamiltonien de Pauli contient d'abord une somme de termes $\sum_i a(\vec{r}_i)\,\vec{l}_i.\vec{s}_i$, en généralisation immédiate de ce qui a été obtenu pour l'unique électron de l'atome d'hydrogène (la fonction $a(r)$ est égale à la fonction notée $A(r)$ en (20.20)). En outre, il apparaît, tout aussi naturellement, des termes en $\vec{s}_i.\vec{l}_j$, appelés *couplages spin - autre orbite*, qui représentent l'interaction entre le moment magnétique du i^{e} électron avec la boucle de courant due au mouvement du j^{e} électron. Enfin, on trouve évidemment l'interaction attendue en $\vec{s}_i.\vec{s}_j$, qui est simplement l'interaction dipolaire entre deux moments magnétiques. Au total, l'opérateur responsable de la structure fine d'un atome complexe a l'allure suivante :

$$\boxed{V_{\text{magn}} = \sum_{i=1}^{N} a(\vec{r}_i)\,\vec{l}_i.\vec{s}_i \;+ \sum_{i,j=1(i\neq j)}^{N} \left[b(\vec{r}_i,\vec{r}_j)\vec{l}_i.\vec{s}_j + f(\vec{r}_i,\vec{r}_j)\,\vec{s}_i.\vec{s}_j\right]} \tag{27.16}$$

La complexité de cet opérateur tient aussi en ce qu'il fait formellement sortir du modèle à particules indépendantes. Pour les éléments légers, ce caractère est cependant moins problématique que la répulsion électrostatique entre les électrons, puisque tous ces termes sont d'ordre α^2 relativement à l'énergie de l'atome en l'absence de structure fine, et sont de toute façon raisonnablement traités par perturbation. L'interaction spin - spin peut toutefois jouer un rôle qualititativement important pour l'ordre des niveaux, même pour un élément léger (l'hélium par exemple).

S'agissant ici de donner les idées principales, on s'en tiendra dans la suite à ne retenir que le premier terme de l'expression (27.16). À des exceptions notables près (dont l'hélium, précisément), cette simplification donne d'ailleurs une description quantitativement acceptable. On admettra donc désormais que la structure fine d'un atome complexe est assez bien prise en compte par l'opérateur simplifié :

$$\boxed{W_{\text{so}} \stackrel{\text{déf}}{=} \sum_{i=1}^{N} a(\vec{r}_i)\,\vec{l}_i.\vec{s}_i} \tag{27.17}$$

Dans cette approximation, le Hamiltonien à prendre en compte est $H_{\text{sf}} \stackrel{\text{déf}}{=} H_{\text{e}} + W_{\text{so}}$, où H_{e} est donné en (27.1).

Cela étant admis, il est clair que les moments cinétiques totaux orbital et de spin, $\vec{L} = \sum_i \vec{l}_i$ et $\vec{S} = \sum_i \vec{s}_i$, ne sont plus des constantes du mouvement. Par exemple :

$$\begin{aligned}\left[\vec{L}^2, W_{\text{so}}\right] &= \sum_i \sum_{j,k} \sum_{u,v=x,y,z} a(\vec{r}_i)\left[l_{uj}l_{uk},\, l_{vi}s_{vi}\right] \\ &= \mathrm{i}\hbar \sum_i \sum_{j,k} \sum_{u,v} a(\vec{r}_i)\varepsilon_{uvw}(\delta_{ki}l_{uj}l_{wk} + \delta_{ji}l_{wj}l_{uk})s_{vi} \;;\end{aligned} \tag{27.18}$$

on a utilisé $[J_u,\, J_v] = \mathrm{i}\hbar\varepsilon_{uvw}J_w$, où ε_{uvw} est le tenseur antisymétrique[10]. Il vient ainsi :

$$[\vec{L}^2, W_{\text{so}}] = \mathrm{i}\hbar\sum_i \sum_{u,v=x,\,y,\,z} a(\vec{r}_i)s_{vi}\varepsilon_{uvw}\sum_j(l_{uj}l_{wi} + l_{wi}l_{uj}) \ . \tag{27.19}$$

La parenthèse à droite est toujours symétrique dans l'échange $u \leftrightarrow v$: si $i \neq j$, les opérateurs commutent, si $i = j$ la symétrie saute aux yeux. Comme ε_{uvw} est antisymétrique en u et v, la somme sur u donne zéro, d'où $[\vec{L}^2, W_{\text{so}}] = 0$; de la même façon, on voit que $[\vec{S}^2, W_{\text{so}}] = 0$. En revanche, il est clair que L_u $(u = x,\, y,\, z)$ ne commute pas avec W_{so} :

$$[L_u, W_{\text{so}}] = \sum_i\sum_j a(\vec{r}_i)[l_{uj},\, l_{vi}s_{vi}] = \sum_i a(\vec{r}_i)[l_{ui},\, l_{vi}]s_{vi} =$$
$$\mathrm{i}\hbar\sum_i \varepsilon_{uvw}a(\vec{r}_i)l_{wi}s_{vi} \neq 0 \ , \tag{27.20}$$

et de même pour $[S_u,\, W_{\text{so}}]$. Ainsi, les *vecteurs* $\vec{L}$ et $\vec{S}$ ne sont plus des constantes du mouvement, mais leurs modules le sont toujours – sans surprise d'ailleurs, puisque le même phénomène se produit déjà pour un électron (voir chapitre 20). On peut donc encore affecter les nombres quantiques L et S aux états de structure fine, mais plus les M_L et M_S, tant que la description simplifiée par W_{so} est adéquate. En raison de la simplification due au passage de V_{magn} à W_{so}, tout ceci est presque trivialement la généralisation à N électrons de ce qui a été fait dans le chapitre 20 pour un électron.

À l'inverse, $\vec{J} = \vec{L} + \vec{S}$ est une constante du mouvement, c'est le moment cinétique total d'un système isolé (si on fait tourner tous les $\vec{l}_i$ et tous les $\vec{s}_i$, la somme des produits scalaires dans W_{so} est évidemment invariante – l'invariance vaut d'ailleurs tout autant pour V_{magn}) ; cette propriété se vérifie d'ailleurs explicitement en calculant le commutateur de $\vec{J}$ avec H_{sf}. Il en résulte que les vecteurs propres pertinents changent de nature ; il faut maintenant considérer[11] les $|\tau JM_JLS\rangle$ au lieu des $(2S+1)(2L+1)$ vecteurs $|\tau LSM_LM_S\rangle$. On passe des uns aux autres par des combinaisons linéaires contenant les coefficients de Clebsch-Gordan, la transformation se faisant à l'intérieur du sous-espace L et S fixés ; les valeurs possibles de J sont :

$$J = L+S,\, L+S-1,\, \ldots,\, |L-S| \ . \tag{27.21}$$

Une conséquence importante est que tout élément de matrice du Hamiltonien entre deux états du genre $|\tau JM_JLS\rangle$ est nul si l'un au moins des nombres quantiques de moment angulaire est différent du *bra* au *ket* :

$$\langle\tau' J'M'_{J'}L'S'|H_{\text{sf}}|\tau JM_JLS\rangle \propto \delta_{JJ'}\delta_{M_JM'_{J'}}\delta_{LL'}\delta_{SS'} \ . \tag{27.22}$$

Comme toujours, les considérations de symétrie permettent de simplifier grandement le problème.

[10] $\varepsilon_{uvw} = -\varepsilon_{vuw}$, $\varepsilon_{uvw} = +1$ si uvw se déduit de xyz par une permutation circulaire, -1 dans le cas contraire

[11] τ désigne l'ensemble des autres nombres quantiques nécessaires pour spécifier complètement l'état. Essentiellement, ce sont ceux qui sont associés à l'énergie propre du Hamiltonien électrostatique (sans structure fine).

Les écarts de structure fine varient comme α^4, et sont donc une petite correction[12] pour les niveaux de type électrostatique (strictement non-relativistes, dont l'énergie est en α^2), qu'ils soient traités à l'intérieur de l'approximation des électrons indépendants, ou que la répulsion de Coulomb entre ceux-ci soit partiellement prise en compte par un champ central effectif. Il est donc légitime de la traiter par une technique de perturbation pour un niveau dégénéré. L'approximation consiste donc ici à négliger les éléments de matrice du type[13] $\langle \tau JM_JLS|H_{\text{sf}}|\tau' JM_JLS\rangle$, $\tau' \neq \tau$; cette procédure est valide tant que l'inégalité $|\langle \tau JM_JLS|W_{\text{so}}|\tau' JM_JLS\rangle| \ll |E_\tau - E_{\tau'}|$ est satisfaite.

Dans ce cadre, les niveaux de structure fine perturbés sont obtenus par diagonalisation de W_{so} projeté dans le sous-espace dégénéré engendré par les $(2L+1)(2S+1)$ vecteurs $|\tau LSM_LM_S\rangle$ (terme spectral), ou les $|\tau JM_JLS\rangle$ correspondants (évidemment en nombre égal). La première base fournit une matrice non-diagonale pour l'opérateur W_{so} projeté ; au contraire, sur la seconde base, W_{so} possède une matrice diagonale (voir (27.22)) : il est donc plus simple et plus direct de se placer d'emblée sur la base $|\tau JM_JLS\rangle$; alors, la lecture des éléments diagonaux fournira les corrections d'énergie au premier ordre (comme L et S sont fixés une fois pour toutes, on note désormais éventuellement $|\tau JM_JLS\rangle \equiv |\tau JM_J\rangle$ pour simplifier). Chaque valeur de J spécifie un *multiplet* : sous l'effet de l'interaction spin-orbite, un même terme spectral se décompose en plusieurs multiplets, ceux-ci constituant la *structure fine* de la raie fortement dégénérée associée au terme spectral considéré.

Pour calculer ces éléments diagonaux, il suffit d'invoquer à plusieurs reprises le théorème de Wigner-Eckart, appliqué au cas des opérateurs vectoriels (voir section 18.6). Selon (18.251), pour tout opérateur vectoriel $\vec{V}$, on a la relation :

$$\langle \tau JM_J|\vec{V}|\tau JM'_J\rangle = \frac{\langle \vec{J}.\vec{V}\rangle}{\hbar^2 J(J+1)}\,\langle \tau JM_J|\vec{J}|\tau JM'_J\rangle \ . \tag{27.23}$$

On choisit d'abord $\vec{J} = \vec{L}$ et $\vec{V} = \vec{l_i}$ dans (27.23) ; ceci permet de montrer que, sur la base $|\tau LSM_LM_S\rangle$, $\vec{l_i}$ et $\vec{L}$ ont des éléments de matrice proportionnels ; la même opération faite avec $\vec{J} = \vec{S}$ et $\vec{V} = \vec{s_i}$ établit que $\vec{s_i} \propto \vec{S}$. Au total, sur cette base, les éléments de matrice de W_{so} et de $\vec{L}.\vec{S}$ sont proportionnels. Cette propriété est évidemment préservée quand on passe aux $|\tau JM_JLS\rangle$ (la transformation (unitaire) se fait à L, S fixés). En définitive, pour un même terme spectral $^{2S+1}\mathcal{F}$ (correspondant précisément à L, S donnés), il est

[12]Pour un élément léger (l'hydrogène par exemple), les variations dues à la structure fine sont de l'ordre de 10 à 100 GHz – nettement inférieures à un élargissement Doppler à temérature ambiante $(k_{\text{B}}T/h)_{300\,\text{K}} \simeq 6\,250$ GHz). Ces corrections, étant d'origine relativiste, sont d'autant plus importantes que l'atome est lourd : l'énergie de l'atome variant en gros comme Z^2, le théorème du Viriel permet d'affirmer, au moins semi-quantitativement, que la vitesse quadratique moyenne d'un électron proche du noyau augmente elle aussi comme Z^2. Pour un atome lourd, les électrons internes vont très vite (ils voient une grande charge nucléaire, et doivent donc tourner très vite pour rester en équilibre) – leurs énergies atteignent donc rapidement quelques dizaines de keV, ce qui n'est pas si petit devant 511 keV ; au contraire, les électrons périphériques voient une charge nucléaire fortement écrantée par les électrons internes, de sorte que leurs énergies restent dans le domaine de l'eV, échelle typique de la Chimie. Les corrections de structure fine sont décroissantes avec le nombre quantique principal.

[13]Ces éléments de matrice n'ont aucune raison d'être nuls, puisqu'ils ont les mêmes nombres quantiques pour tous les moments cinétiques.

équivalent, à un facteur près, de manipuler W_{so} ou $\vec{L}.\vec{S}$, ce que l'on résume par :

$$\boxed{W_{\text{so}} \longrightarrow A(\tau,\, L,\, S)\,\vec{L}.\vec{S} \quad \text{pour un multiplet donné}} \tag{27.24}$$

où $A(\tau,\, L,\, S)$ est une constante caractéristique du terme spectral, ou du multiplet, considéré[14]. L'opérateur $\vec{L}.\vec{S}$ est visiblement de trace nulle (elle l'est trivialement sur la base $|\tau LSM_LM_S\rangle$, et la trace est un invariant) : la somme des corrections d'énergie pondérée par les dégénérescences est donc nulle. Finalement, utilisant $\vec{L}.\vec{S} = \frac{1}{2}(\vec{J}^{\,2} - \vec{L}^{\,2} - \vec{S}^{\,2})$, on obtient l'expression des corrections d'énergie produisant la structure fine :

$$\boxed{\Delta E_{\text{so}} = \frac{\hbar^2}{2}\,A(\tau,\, L,\, S)\,[J(J+1) - L(L+1) - S(S+1)]} \tag{27.25}$$

Ce résultat s'appelle la *règle des intervalles de Landé*, qui stipule [210] que la variation d'énergie est de la forme $\frac{1}{2}AJ(J+1) + \text{C}^{\text{ste}}$. En l'absence de champ magnétique, tout multiplet $|\tau JM_JLS\rangle$ reste dégénéré $2J+1$ fois. Par exemple, pour un terme spectral du type ^2P, J prend les deux valeurs $1 \pm \frac{1}{2}$, soit $\frac{1}{2}$ et $\frac{3}{2}$; l'usage est de compléter la notation du terme spectral en indiquant la valeur de J en indice : dans le cas présent, on obtient les deux multiplets $^2\text{P}_{\frac{1}{2}}$ et $^2\text{P}_{\frac{3}{2}}$. Les dégénérescences respectives sont 2 et 4, leur somme vaut 6 et est bien égale à la dégénérescence ($= 2 \times 3$) du terme spectral ^2P qui a donné naissance aux deux multiplets différant par leur J.

27.4 Structure hyperfine

Le noyau de l'atome est formé de nucléons[15], protons et neutrons, qui sont tous des fermions de spin $S = 1/2$. L'addition de tous ces moments cinétiques constitue le spin du noyau, usuellement désigné par $\vec{I}$. À ce moment cinétique est associé un moment magnétique qui peut se coupler d'une part aux moments magnétiques des électrons suivant une interaction de type dipôle - dipôle, d'autre part au champ magnétique résultant des petites boucles de courant que constituent les électrons en orbite – pour employer un langage imagé. Une fois introduit le spin nucléaire, le moment cinétique de l'atome est la somme $\vec{F} \stackrel{\text{déf}}{=} \vec{I} + \vec{J} = \vec{I} + \vec{L} + \vec{S}$, et c'est vis-à-vis de ce vecteur total que l'atome est maintenant globalement invariant par rotation. L'interaction de $\vec{I}$ avec $\vec{J}$ donne lieu à de nouvelles corrections pour l'énergie de l'atome, qui sont le plus souvent d'un ordre de grandeur plus petites que celles dues à la structure fine, d'où la terminologie *structure hyperfine*, et sont particulièrement importantes pour l'état fondamental (qui implique des électrons près du noyau).

Par exemple, c'est la structure hyperfine du fondamental de l'hydrogène (neutre) qui donne lieu à la célèbre raie à 21 cm, correspondant à un écart de niveaux de l'ordre

[14] Cette substitution définit le couplage dit LS, ou *de Russel - Saunders*, schéma réservé aux éléments légers. Pour les atomes lourds, l'interaction dominante est entre chaque $\vec{l}_i$ et $\vec{s}_i$, ce qui impose de travailler en considérant d'abord les $\vec{j}_i = \vec{l}_i + \vec{s}_i$, puis de coupler les $\vec{j}_i$ entre eux (couplage $j - j$).

[15] Point besoin d'aller jusqu'aux *quarks*...

de 6×10^{-6} eV, soit environ[16] 1,4 Ghz ; c'est cette raie qui permet de partir à la chasse à l'hydrogène dans les milieux astrophysiques : quoique la probabilité de transition est très faible, la quantité (astronomique !) d'hydrogène rend cette raie observable. On se souvient aussi que l'observation de l'effet Zénon par Itano et al. [98] implique la structure hyperfine du fondamental de l'ion Be^+ (voir fig. 21.2).

La petitesse de l'interaction hyperfine tient essentiellement à la grande masse des nucléons, comparée à celle de l'électron. Dans le cas simple de l'hydrogène, le noyau se réduit à un proton, de masse $M_\text{p} \simeq 1\,840\,m$; le facteur anormal du proton est $g_\text{p} \simeq 5,586$, de sorte que le moment magnétique du proton μ_p, égal à $\gamma_\text{p}\frac{|e|}{2M_\text{p}}$, vaut environ $1,5 \times 10^{-3}$ fois celui de l'électron. Ainsi, le plus souvent (mais pas toujours) la structure hyperfine donne des corrections encore plus petites que celles dues à la structure fine. Par ailleurs, et pour les mêmes raisons tenant à leur origine relativiste, elles décroissent quand le nombre quantique principal augmente.

L'écriture détaillée de l'opérateur d'interaction fait apparaître différents termes. Le premier d'entre eux est du type $\vec{L}.\vec{I}$, représentant l'interaction spin nucléaire - orbite des électrons, et peut s'obtenir exactement comme pour l'interaction spin - orbite de l'électron (voir p. 849) – sans le petit problème lié à la précession de Thomas puisque l'on se place d'emblée dans le repère du noyau. Le couplage correspondant est de la forme $A(r)\vec{L}.\vec{I}$, où la fonction $A(r)$ est celle qui figure déjà en (20.20), où il faut juste remplacer g_e par g_p et m^2 par mM_p. Un autre terme d'interaction représente le couplage entre le spin électronique et le champ magnétique créé par le spin nucléaire[17], qui a l'allure classique – à un terme de contact près traduisant la fermeture des lignes de champ le long d'une boucle dont la dimension tend vers zéro[18]. Ce dernier terme ne joue que pour les états électroniques de type S, où la fonction d'onde est non-nulle en $r = 0$. Au total, l'expression complète de l'opérateur de structure hyperfine, pour l'hydrogène neutre, est la suivante[19] :

$$\boxed{V_\text{shf} = \frac{g_\text{e}g_\text{p}e'^2}{4mM_\text{p}c^2}\left\{\frac{1}{r^3}\vec{L}.\vec{I} - \left[\frac{1}{r^3}\vec{I}.\vec{S} - \frac{3}{r^5}(\vec{I}.\vec{r})\,(\vec{S}.\vec{r})\right] + \frac{8\pi}{3}\vec{I}.\vec{S}\,\delta(\vec{r})\right\}} \tag{27.26}$$

Compte tenu de la petitesse des termes de couplage – y compris comparés à ceux donnant la structure fine (le plus souvent) –, les corrections d'énergie peuvent être en toute sécurité calculées avec la théorie des perturbations. En particulier, le couplage

[16]On connaît la valeur de la fréquence à une très grande précision : $\nu_\text{21cm} = 1,420\,405\,751\,766\,7$ Ghz. Dans les conditions terrestres, la raie est quasi inobservable en émission spontanée, mais la fréquence a pu être mesurée par émission stimulée.

[17]Se souvenir que spin = moment magnétique.

[18]Classiquement, ce terme n'apparaît pas (il fut d'ailleurs oublié au début des applications de la théorie quantique), puisque le champ dipolaire est calculé loin de la boucle de courant, plus précisément à une distance grande devant le rayon de celle-ci ; de façon équivalente, à une distance donnée, le champ à grande distance est obtenu en faisant tendre vers zéro l'extension spatiale de la boucle. Ici, la longueur qui tend vers zéro (ce qui veut dire qu'elle est aussi petite que l'on veut) est le rayon du proton.

Le même phénomène se produit pour l'interaction spin - spin mentionnée à propos de la structure fine (voir éq. (27.16), termes en $\vec{s}_i.\vec{s}_j$) et provoque l'apparition de termes de contact $\delta(\vec{r}_i - \vec{r}_j)$.

[19]Pour plus de détails, voir [89].

spin - orbite donne lieu à une règle des intervalles de Landé d'un nouveau genre, avec une correction proportionnelle à $[F(F+1) - J(J+1) - I(I+1)]$; le calcul précis des différentes corrections montre que celles-ci sont d'ordre de grandeur $\frac{m}{M_{\rm p}}\alpha^4\, mc^2$: elles sont bien typiquement environ $\frac{m}{M_{\rm p}}$ fois plus petites que les corrections de structure fine.

Pour l'hydrogène dans l'état fondamental, $L = 0$ et $S = 1/2$, de sorte que J prend l'unique valeur $J = 1/2$ (le fondamental n'a pas de structure fine). Dès que l'on introduit le couplage hyperfin, le spin total $\vec{F} = \vec{I} + \vec{J}$ est la somme de deux moments cinétiques 1/2 et le nombre F peut prendre les deux valeurs $F = 0$ (singulet) et $F = 1$ (triplet) : l'état fondamental se *dédouble*, la différence d'énergie pouvant se calculer par perturbation au premier ordre. La symétrie de l'état $L = 0$ entraîne que chacun des deux premiers termes donne une correction nulle[20], de sorte que toute la correction vient du terme de contact – d'où son importance. La fonction de Dirac introduit[21] $|R_{10}(0)Y_{00}|^2 = \frac{1}{\pi a_0^3}(1 + m/M_{\rm p})^{-3}$, de sorte que la valeur moyenne du terme de contact dans le multiplet ${}^2S_{1/2}$ est :

$$\frac{2g_{\rm e}g_{\rm p}\alpha^4 m^2c^2}{3\hbar^2 M_{\rm p}(1+\frac{m}{M_{\rm p}})^3}\,\langle \vec{I}.\vec{S}\,\rangle\ . \tag{27.27}$$

Pour trouver $\langle \vec{I}.\vec{S}\,\rangle$, on écrit $\vec{I}.\vec{S} = \frac{1}{2}(\vec{F}^{\,2} - \vec{I}^{\,2} - \vec{S}^{\,2})$, qui vaut $\frac{\hbar^2}{4}$ dans l'état triplet $F = 1$ et $-\frac{3\hbar^2}{4}$ dans l'état singulet. En définitive, on trouve :

$$E_{{}^2{\rm S}_{1/2}}(F=1) - E_{{}^2{\rm S}_{1/2}}(F=0) = \frac{2}{3}\,g_{\rm e}g_{\rm p}\alpha^4\,\frac{m}{M_{\rm p}(1+\frac{m}{M_{\rm p}})^3}\,mc^2\ . \tag{27.28}$$

Numériquement, avec $g_{\rm e} = 2(1 + \frac{\alpha}{2\pi} + ...)$, on trouve :

$$\boxed{E_{{}^2{\rm S}_{1/2}}(F=1) - E_{{}^2{\rm S}_{1/2}}(F=0) = 5{,}874... \times 10^{-6}\,{\rm eV} = h \times 1{,}420...{\rm Ghz}} \tag{27.29}$$

C'est la fameuse raie 21 cm, manifestation d'une transition dipolaire magnétique (M1). Comme attendu, le *splitting* hyperfin est beaucoup plus petit pour les états excités ; par exemple, pour l'état $n = 2$, il vaut environ 24 MHz pour le niveau ${}^2{\rm P}_{3/2}$, 178 MHz pour ${}^2{\rm S}_{1/2}$ et 59 MHz pour ${}^2{\rm P}_{1/2}$. On sait que la dégénérescence du couple (${}^2{\rm P}_{1/2}$, ${}^2{\rm S}_{1/2}$) est levée par des corrections radiatives (déplacement de Lamb, voisin de 1 060 MHz, voir p. 898) ; pour ces niveaux excités, la structure hyperfine donne des écarts d'énergie nettement plus petits.

Pour un atome à plusieurs électrons, l'opérateur de structure hyperfine devient une somme d'opérateurs à un seul électron, puisque le couplage a lieu entre le noyau et les électrons, non pas entre les électrons. Procédant comme ci-dessus pour l'hydrogène, il convient ensuite d'ajouter ces différentes contributions, compte tenu de la symétrie de moment cinétique des différentes couches concernées. Un traitement détaillé pour l'hélium peut être trouvé dans [89], § 44.

[20]En ce qui concerne le terme dipôle - dipôle, c'est plutôt une bonne chose... puisque la valeur moyenne de r^{-3} dans un état sphérique diverge logarithmiquement à l'origine.

[21]Pour la précision, il est important ici de ne pas oublier que la première orbite de Bohr a_0 est en toute rigueur définie avec la masse *réduite* μ ; toujours pour la précision, il ne faut évidemment pas remplacer $g_{\rm e}$ par sa valeur approchée $g_{\rm e} \simeq 2$.

27.5 Théorème du Viriel

Dans l'impossibilité de résoudre exactement le problème à N- corps, force est de recourir à des méthodes approchées dont il faut pouvoir contrôler la qualité. Celle-ci peut s'apprécier de diverses façons et par référence au but poursuivi ; elle peut également être mesurée par confrontation avec des propriétés générales que possède la solution exacte inconnue. Ces propriétés peuvent être des symétries requises, ou la satisfaction de relations fondamentales. Un exemple est fourni par le théorème du Viriel, bien connu en Mécanique classique et qui se généralise au domaine quantique ; la discussion de ce théorème permet d'ailleurs de réaliser que certains garde-fous peuvent se révéler illusoires.

Dans ce qui suit[22], on considère uniquement les interactions électrostatiques de Coulomb, qui sont des fonctions homogènes de degré -1 ; pour tout terme de Coulomb, $v(r)$, on a :

$$v(\lambda r) = \lambda^{-1}\, v(r) \ , \tag{27.30}$$

où λ est un facteur d'échelle de contraction ou de dilatation des longueurs (*scaling*). En particulier, si on note $\mathcal{V}(\vec{r}_1,\, \vec{r}_2,\, \ldots,\, \vec{r}_N)$ l'ensemble des termes coulombiens (attraction du noyau et répulsion entre électrons), on a :

$$\mathcal{V}(\lambda\vec{r}_1, \lambda, \vec{r}_2,\, \ldots,\, \lambda\vec{r}_N) = \lambda^{-1}\, \mathcal{V}(\vec{r}_1,\, \vec{r}_2,\, \ldots,\, \vec{r}_N) \ . \tag{27.31}$$

L'énergie potentielle électrostatique totale est ainsi une fonction homogène de degré -1. Par ailleurs, en représention-q, tout terme cinétique est du genre $-\frac{\hbar^2}{2m}(\vec{\nabla}_{\vec{r}})^2$: c'est donc aussi une fonction homogène des coordonnées, mais de degré -2.

$\Psi(\vec{r}_1,\, \vec{r}_2,\, \ldots,\, \vec{r}_N)$ désignant une fonction propre exacte normalisée[23], définissons une autre fonction déduite de celle-ci[24] par un *scaling* :

$$\Psi_\lambda(\vec{r}_1,\, \vec{r}_2,\, \ldots,\, \vec{r}_N) \overset{\text{déf}}{=} \lambda^{\frac{3N}{2}}\, \Psi(\lambda\vec{r}_1,\, \lambda\vec{r}_2,\, \ldots,\, \lambda\vec{r}_N) \ . \tag{27.32}$$

Le facteur $\lambda^{\frac{3N}{2}}$ assure que Ψ_λ est normalisée au même titre que Ψ. Calculons maintenant les valeurs moyennes de l'énergie cinétique et de l'énergie potentielle dans l'état (fictif) Ψ_λ. En changeant les variables muettes dans les intégrales, on établit facilement :

$$\langle T\rangle(\lambda) \overset{\text{déf}}{=} \langle\Psi_\lambda|T|\Psi_\lambda\rangle = \lambda^2\, \langle\Psi|T|\Psi\rangle \equiv \lambda^2\, \langle T\rangle \ , \tag{27.33}$$

$$\langle\mathcal{V}\rangle(\lambda) \overset{\text{déf}}{=} \langle\Psi_\lambda|\mathcal{V}|\Psi_\lambda\rangle = \lambda\, \langle\Psi|\mathcal{V}|\Psi\rangle \equiv \lambda\, \langle\mathcal{V}\rangle \ . \tag{27.34}$$

Maintenant, par le Principe variationnel, on peut affirmer que dans la famille de fonctions $\{\Psi_\lambda\}_\lambda$ il existe une fonction donnant sa plus petite valeur à l'énergie $E(\lambda)$ donnée par :

$$E(\lambda) = \langle T\rangle(\lambda) + \langle\mathcal{V}\rangle(\lambda) \ . \tag{27.35}$$

[22]L'argument suivant est dû à Fock.

[23]En pratique, il s'agira de l'état fondamental puisque la suite de l'argument s'appuie sur le Principe variationnel.

[24]Ψ coïncide formellement avec $\Psi_{\lambda=1}$.

Compe tenu de (27.33) et (27.34), la condition de minimisation de $E(\lambda)$ est :

$$2\lambda \langle T\rangle + \langle \mathcal{V}\rangle = 0 \ . \qquad (27.36)$$

Comme la famille variationnelle contient la solution exacte Ψ, qui n'est autre que $\Psi_{\lambda=1}$, la relation (27.36) est vraie en $\lambda = 1$. On en déduit que, pour la fonction exacte, le théorème du Viriel pour le champ Coulombien s'énonce :

$$\boxed{2\langle T\rangle = -\langle \mathcal{V}\rangle} \qquad (27.37)$$

On peut donc vouloir exiger que toute fonction approchée satisfasse ce théorème, en gage de sa qualité. Ce critère n'est pas forcément sûr. En effet, supposons maintenant que la fonction Ψ ci-dessus *n'est pas* la fonction exacte, avec laquelle, toujours selon la définition (27.32), on peut définir une famille variationnelle de fonctions Ψ_λ. La relation de minimisation de l'énergie, qui donne maintenant la valeur de λ définissant la meilleure fonction au sens du Principe Variationnel, ne fournit plus $\lambda = 1$ (selon (27.36), la valeur optimale de λ est donnée par le rapport $-\frac{\langle \mathcal{V}\rangle}{2\langle T\rangle}$) ; cependant, la même relation de minimisation s'écrit aussi :

$$2\lambda^2 \langle T\rangle + \lambda\langle \mathcal{V}\rangle = 0 \ , \qquad (27.38)$$

soit, compte tenu de (27.33) et (27.34) :

$$2\langle T\rangle(\lambda) = -\langle \mathcal{V}\rangle(\lambda) \ . \qquad (27.39)$$

Autrement dit, disposant d'une fonction approchée quelconque, un simple *scaling* permet de construire une autre fonction approchée satisfaisant le théorème du Viriel : il suffit de faire une nouvelle variation en minimisant l'énergie par rapport au facteur d'échelle. En conclusion, la satisfaction du théorème du Viriel n'est pas, en soi, un signe de la qualité d'une fonction d'onde approchée.

En outre, la portée de ce théorème est souvent limitée en pratique du fait que le potentiel *nu* constitue une idéalisation qui n'est pas toujours acceptable. En particulier, en présence d'écrantage par les électrons internes, le potentiel ressenti par les électrons de valence est à grande distance plutôt du genre Yukawa et invalide de ce fait les hypothèses conduisant à une relation comme celle qu'exprime le théorème du Viriel.

27.6 Au-delà de l'approximation à électrons indépendants

Le modèle à particules indépendantes doit être considéré comme une description minimale qu'il faut s'empresser de dépasser pour rendre compte des propriétés les plus élémentaires des systèmes de particules couplées les unes aux autres. Il existe une immense variété de schémas (approximatifs) permettant d'incorporer les interactions, dont le choix et les les détails de procédure dépendent en grande partie du système étudié et des propriétés

auxquelles on s'intéresse plus particulièrement : on n'utilise pas les mêmes méthodes selon qu'il s'agit de calculer avec précision une fréquence de transition atomique, la réactivité d'une molécule ou la structure de bande d'un cristal.

En délaissant toutes les approximations propres aux systèmes à très grand nombre de particules utilisées, par exemple, en physique de la matière condensée, il y a en gros deux grandes familles d'approximations. L'une d'entre elles consiste à utiliser des techniques de type perturbatif, l'autre à s'appuyer sur le Principe variationnel, étant entendu qu'il est toujours possible d'allier les deux façons de faire. Les équations de Hartree - Fock, données dans le chapitre 23 ont été déduites ce Principe. Une autre approche, plus récente, s'appuie sur le théorème de Hohenberg et Kohn [176] et a engendré toutes les méthodes de *fonctionnelle de la densité.*

Des exemples très simples d'application de la méthode variationnelle ont déjà été donnés (chapitre 23), à titre d'illustrations élémentaires. Dans le contexte actuel, il vaut la peine de revenir sur cette méthode en traitant un cas non trivial, mais qui reste simple, susceptible d'ailleurs de recevoir une interprétation physique évidente, à savoir l'atome d'hélium.

Comme on l'a dit, le choix de la famille de fonctions d'essai doit être inspiré par des considérations physiques, complétées évidemment par des prérequis pour la fonction cherchée, ses symétries (sous toute forme) en particulier. S'agissant de traiter l'état fondamental, qui est singulet, il convient donc de prendre une fonction d'essai symétrique dans l'échange de ses arguments. Dès lors, le choix qui s'impose en premier est un simple produit d'une même fonction $\psi(\vec{r})$, prise une fois en $\vec{r}_1$ et l'autre en $\vec{r}_2$. De surcroît, sachant que le fondamental est à symétrie sphérique ($L = 0$), la fonction ψ ne dépend que de $\|\vec{r}\| \equiv r$. Enfin, il est naturel de s'inspirer de la forme des fonctions hydrogénoïdes, ce qui suggère de prendre $\psi(r) \propto \mathrm{e}^{-Zr/a_0}$. Si l'on fixait Z à la valeur $Z = 2$, tout serait déterminé ; l'idée est alors de laisser *flotter* cette quantité, qui se trouve *de facto* parée du statut de paramètre variationnel. On va donc calculer l'énergie avec une telle fonction d'essai, obtenant une *fonction* $E(Z)$, dont le minimum (espérant qu'il existe), survenant pour une valeur Z_{opt}, définira la meilleure fonction de ce type au sens de la méthode variationnelle. Pour des raisons physiques, on s'attend à trouver que Z_{opt} est un peu inférieur à 2 ; l'image physique sous-jacente est l'écrantage de la charge nucléaire par un électron vis-à-vis de l'autre.

Soit donc la fonction normalisée :

$$\psi_Z(r_1,\, r_2) = \frac{Z^3}{\pi a_0^3}\, \mathrm{e}^{-Z\frac{r_1}{a_0}}\, \mathrm{e}^{-Z\frac{r_2}{a_0}} \quad , \tag{27.40}$$

satisfaisant :

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |\psi_Z(r_1,\, r_2)|^2\, \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 = 1 \quad . \tag{27.41}$$

Il faut maintenant calculer la fonctionnelle énergie $E(Z)$ pour le Hamiltonien[25] :

$$H = \frac{\vec{p}_1^{\,2}}{2m} + \frac{\vec{p}_2^{\,2}}{2m} - \frac{2e'^2}{r_1} - \frac{2e'^2}{r_2} + \frac{e'^2}{r_{12}} \ ; \tag{27.42}$$

la fonction ψ étant normalisée, on a simplement :

$$E(Z) = \langle\psi_Z|\frac{\vec{p}_1^{\,2}}{2m} + \frac{\vec{p}_2^{\,2}}{2m} - \frac{2e'^2}{r_1} - \frac{2e'^2}{r_2} + \frac{e'^2}{r_{12}}|\psi_Z\rangle \ . \tag{27.43}$$

Les contributions à un corps ne présentent aucune difficulté de calcul ; on peut d'ailleurs les obtenir par la relation de *scaling* (27.33) et (27.34) puisque visiblement Z a aussi, ici, le rôle d'un facteur d'échelle. Pour chaque électron $\langle T\rangle(Z=1)$ est égal à la valeur moyenne de l'énergie cinétique de l'électron de l'atome d'hydrogène[26], soit $(-E_1) = \frac{e'^2}{2a_0}$: l'énergie cinétique des deux électrons est donc $-2Z^2E_1$. De la même façon, l'énergie potentielle attractive pour les deux électrons et pour $Z=1$ vaut $2\frac{-2e'^2}{a_0} = 8E_1$, d'où $8ZE_1$ pour Z quelconque. Les termes à un électron donnent donc au total $2E_1(4Z - Z^2)$ (qui vaut bien $2\times(2)^2E_1$ si $Z=2$). Le terme à deux corps peut se calculer en développant $\frac{e'^2}{r_{12}}$ en polynômes de Legendre, éq. (18.114), ou en harmoniques sphériques, (18.123) (pour plus de détails, voir [53], p. 258, et section 27.7 ci-après, notamment éq. (27.65)) ; il vaut :

$$\langle\psi_Z\psi_Z|\frac{e'^2}{r_{12}}|\psi_Z\psi_Z\rangle = \frac{5}{8}\frac{Ze'^2}{a_0} \equiv -\frac{5Z}{4}E_1 \ . \tag{27.44}$$

Au total, on trouve :

$$E(Z) = 2E_1(4Z - Z^2 - \frac{5Z}{8}) = 2E_1(\frac{27}{8}Z - Z^2) \ ; \tag{27.45}$$

les énergies totales, cinétique et potentielle, sont bien proportionnelles respectivement à Z^2 et Z, conformément aux relations de *scaling* démontrées dans la section 27.5, et en raison ici, avec le choix (27.40), de la nature particulière du paramètre variationnel Z.

La fonction (27.45) présente bien un minimum, survenant pour une certaine valeur de Z :

$$Z_{\text{opt}} = \frac{27}{16} \simeq 1,69 \ , \tag{27.46}$$

valeur un peu inférieure à 2 comme on s'y attendait. L'énergie correspondante est :

$$E(Z_{\text{opt}}) \simeq -2,848\frac{e'^2}{a_0} \geq -2,904\frac{e'^2}{a_0} \equiv E_{\text{exp}} \ . \tag{27.47}$$

L'erreur sur l'énergie est de l'ordre de 2%, ce qui est assez remarquable compte tenu de la rusticité de la méthode[27].

[25] *Attention !* Le principe de la méthode consiste à calculer la valeur moyenne du Hamiltonien *exact* avec une fonction d'essai choisie : ne pas mettre Z – le paramètre variationnel – dans les termes d'attraction électron-noyau (voir (27.43))...

[26] Le théorème du Viriel donne $2\langle T\rangle + \langle V\rangle = 0$; avec $E = \langle T\rangle + \langle V\rangle$, on a $E = -\langle T\rangle = \frac{1}{2}\langle V\rangle$. Pour le fondamental de l'hydrogène, $E_1 = -\frac{me'^4}{2\hbar^2} = -\frac{e'^2}{2a_0}$.

[27] Par comparaison, une méthode ordinaire de perturbation (où Z est fixé à la valeur 2) donne un résultat entaché d'une erreur de l'ordre de 5% : voir plus loin, p. 1245.

Bien évidemment, cette approche peut être raffinée de bien des façons. Un calcul célèbre [211] est dû à Hylleraas en 1929, qui a considéré des fonctions d'essai dépendant explicitement de la distance r_{12} ; les résultats sont bien meilleurs, mais la méthode n'a pas pu être généralisée au cas d'atomes plus complexes : de ce point de vue, la fonction de Hylleraas est plutôt un remarquable objet de musée.

Signalons au passage une autre méthode, dite de *l'interaction de configuration.* Elle consiste à élargir l'espace de recherche et à considérer un espace vectoriel de fonctions associées à *plusieurs* configurations électroniques ayant les bonnes symétries. Par exemple, même pour le fondamental de l'hélium, il est sensé de considérer tous les déterminants issus par exemple des deux configurations $(1s)^2$ et (1s)(2s), à condition de retenir pour cette dernière le seul état singulet ($S = 0$) qui en résulte. On trouve alors[28], par diagonalisation de la matrice du Hamiltonien, que l'état fondamental est une combinaison linéaire des deux singulets, avec toutefois un poids prépondérant ($\simeq 90\%$) sur le singulet construit avec $(1s)^2$. Il est clair que l'intérêt de cette méthode dépend en grande partie de l'outil numérique disponible. En outre, si l'on peut affirmer qu'en augmentant la dimension de l'espace – c'est-à-dire en augmentant le nombre de configurations de bonne symétrie – on ira forcément vers une amélioration de la description, la convergence peut être très lente, en particulier parce que les fonctions hydrogénoïdes *liées* ne sont pas complètes en raison de l'existence des états du *continuum* associés aux états de diffusion ([212], [213]). Cette méthode peut d'ailleurs être utilisée en conjonction avec une optimisation des paramètres définissant les orbitales entrant dans la fonction d'essai. Enfin, l'emploi d'exponentielles de type hydrogénoïdes n'est nullement une nécessité : on utilise souvent des fonctions *gaussiennes*, plus rapides à calculer par ordinateur.

♦ *Remarques*

1. Le calcul effectué ci-desus pour l'hélium avec la fonction (27.40) n'est pas un calcul SCF : rien ne permet d'affirmer que les orbitales SCF sont les fonctions 1s hydrogénoïdes ! Le choix d'exponentielles dans (27.40) restreint l'ensemble des fonctions participant à la variation, alors que, dans le cadre SCF, on n'impose rien *a priori* (en principe du moins). Pour l'hélium, prenant un simple produit d'une même orbitale (donc forcément pour un état singulet), il n'y a qu'une équation SCF, dénuée d'ailleurs du terme d'échange, qui s'écrit :

$$\Big(-\frac{\hbar^2}{2m}\Delta-\frac{2e'^2}{r}\Big)\psi^{\mathrm{SCF}}(\vec{r})+\int_{\mathbb{R}^3}\mathrm{d}^3r'|\psi^{\mathrm{SCF}}(\vec{r}\,')|^2\frac{e'^2}{\|\vec{r}-\vec{r}\,'\|}\,\psi^{\mathrm{SCF}}(\vec{r}) = E^{\mathrm{SCF}}\psi^{\mathrm{SCF}}(\vec{r})\ , \tag{27.48}$$

dont la solution (de plus basse énergie) se trouve numériquement par un algorithme itératif : on part d'une fonction donnée $\psi^{(0)}(\vec{r})$ avec laquelle on calcule l'intégrale, qui définit à cette étape le champ moyen de l'autre électron ; on cherche ensuite la fonction propre de plus basse énergie, $\psi^{(1)}(\vec{r})$, de ce Hamiltonien effectif ; on recommence avec $\psi^{(1)}(\vec{r})$ pour avoir $\psi^{(2)}(\vec{r})$, et ainsi de suite jusqu'à satisfaction d'un critère de convergence préalablement défini, du genre $\int_{\mathbb{R}^3}|\psi^{(n+1)}(\vec{r}) - \psi^{(n)}(\vec{r})|\,\mathrm{d}^3r < \varepsilon$.

[28] Le calcul se fait à la main sans grande difficulté, voir problème 27.8.7.

Il n'existe pas de théorèmes de convergence, d'autant plus que l'équation (27.48) est non-linéaire, ce qui ouvre la voie aux instabilités, aux vraies fausses solutions[29] ou aux brisures de symétrie, de spin notamment.

2. Les fonctions hydrogénoïdes sont paramétrées par le numéro atomique Z, lequel est en fait un *facteur d'échelle* puisque toutes les fonctions radiales sont de la forme $R_{nl}(r) = Z^{3/2} f_{nl}(\frac{Zr}{a_0})$. La dépendance en Z de toute contribution à l'énergie calculée avec de telles fonctions s'en déduit donc immédiatement : pour l'énergie cinétique $\langle T\rangle(Z) = Z^2\langle T\rangle(1)$, et, s'il n'y a que des interactions de Coulomb, $\langle V\rangle(Z) = Z\langle V\rangle(1)$. ♦

27.7 L'atome d'hélium

L'atome d'hélium ($Z = 2$) est l'atome le plus simple[30] contenant déjà la complexité foncière des systèmes polyélectroniques. Une description précise doit évidemment inclure les termes d'interaction spin-orbite donnant lieu à la structure fine mais, pour s'en tenir à l'approche la plus simple, on ne retiendra que les interactions de Coulomb. Il s'agit juste ici de rentrer un peu dans le détail du calcul explicite de quantités typiques apparaissant inévitablement dès que l'on souhaite faire une étude quantitative en Physique atomique, ou en Chimie quantique.

Dans ce cadre, le Hamiltonien a l'expression donnée en (27.42) ou, sous une forme plus concise :

$$H(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2) = [H_1(\vec{r}_1, \vec{p}_1) + H_1(\vec{r}_2, \vec{p}_2)] + \frac{e'^2}{r_{12}} \equiv H_0 + V \ . \tag{27.49}$$

$H_1(\vec{r}, \vec{p}) \stackrel{\text{déf}}{=} \frac{\vec{p}^2}{2m} - 2\frac{e'^2}{r}$ est le Hamiltonien hydrogénoïde pour $Z = 2$. On note au passage que H est bien invariant dans toute permutation, comme il se doit pour des particules identiques, ici les deux électrons.

Le point de départ est à nouveau le modèle à électrons indépendants. Ceci revient à considérer d'abord H_0 défini en (27.49), qui est une somme d'opérateurs à une particule dont on connaît les solutions exactes ; la perturbation $V = \frac{e'^2}{r_{12}}$ étant positive, les vrais niveaux d'énergie de l'hélium seront plus haut en énergie que leurs homologues obtenus en tant qu'états propres de H_0.

[29]On connaît des cas très simples d'équations non-linéaires, dont la solution est accessible analytiquement, et pour lesquelles les schémas itératifs donnent des résultats grossièrement faux (fourniture d'une solution instable alors qu'il existe une solution stable), quand ils ne sont pas tout simplement stupides (du genre : *"Il n'y a pas de solution"*), alors que l'on sait bien qu'elle existe, puisqu'on l'a trouvée *à la main*...

[30]L'hélium a été découvert en 1868 par un astronome français, Pierre Janssen, profitant d'une éclipse de Soleil pour faire l'analyse du spectre solaire, et y découvrant une raie jaune alors inconnue.

Si $\phi_n \equiv \phi_{nlm}$ est une fonction propre hydrogénoïde avec $Z = 2$, d'énergie[31] :

$$E_n = -\frac{m(2e'^2)^2}{2n^2\hbar^2} \equiv -\frac{4}{n^2}E_\mathrm{I} \simeq -\frac{54,42}{n^2}\,\mathrm{eV} \ , \tag{27.50}$$

alors tout produit $\phi_{n_1}\phi_{n_2}$ est fonction propre de H_0 :

$$H_0\,\phi_{n_1}(\vec{r}_1)\phi_{n_2}(\vec{r}_2) = (E_{n_1} + E_{n_2})\,\phi_{n_1}(\vec{r}_1)\phi_{n_2}(\vec{r}_2) \equiv E_{n_1n_2}\phi_{n_1}(\vec{r}_1)\phi_{n_2}(\vec{r}_2) \ . \tag{27.51}$$

L'expression (27.50) montre que l'état fondamental de l'ion He^+ a pour énergie -54,42 eV. Par ailleurs, l'état hydrogénoïde $n = 2, Z = 2$ a pour énergie $-13,61$ eV ; la configuration de l'hélium (1s)(2s) (ou (1s)(2p)) – un seul électron excité – a donc une énergie égale à $-68,03$ eV, qui est au-dessous du fondamental de He^+. Toute excitation du deuxième électron donne un état instable où l'atome l'ionise spontanément, ceci sans compter avec l'effet déstabilisant de la répulsion électronique qui accentue encore les choses.

L'instabilité vient du fait qu'un état comme $\phi_{200}(r_1)\phi_{200}(r_2)$, d'énergie égale à $-27,22$ eV, est en face du *continuum* d'états $\phi_{100}(r_1)\phi_{E00}(r_2)$, d'énergie $-54,42$ eV $+E$ ($E > 0$ variable), représentant un électron dans le fondamental de He^+ et l'autre dans l'état non-lié d'énergie E de cet ion. On est donc typiquement dans la situation d'un état discret en regard d'un continuum, pour lequel la Règle d'or de Fermi donne une évolution irréversible et une instabilité pour l'état discret (voir section 24.4). Pour que ceci se produise, il faut bien sûr un couplage enre l'état discret et le *continuum*, et c'est la répulsion électronique qui s'en charge[32]. Au total, l'atome d'hélium s'ionise spontanément avec une excitation bien inférieure à $4E_\mathrm{I}$, énergie du fondamental hydrogénoïde avec $Z = 2$.

D'ailleurs, les états $\phi_{100}(r_1)\phi_{nlm}(r_2)$ avec n assez grand, sont aussi instables, et pour les mêmes raisons. En définitive, et s'agissant de décrire les états stables de He, seuls comptent les états où l'un des électrons reste au fondamental hydrogénoïde, et où l'autre est dans une orbitale modérément excitée ; cela n'aurait aucun sens de considérer des configurations $\phi_n\phi_{n'}$ arbitraires. Il n'en reste pas moins que des expériences récentes [214] ont pu mettre en évidence le déclin *radiatif* de ces états métastables, qui présentent un grand intérêt sur le plan fondamental.

Les ϕ_n sont des orbitales ; il convient de les multiplier par une fonction de spin ($= \alpha$ ou β) pour former une spin orbitale, qui est évidemment toujours propre de H_0 (H_0, pas plus que H d'ailleurs, ne contient le spin) ; avec le produit des deux spin-orbitales, il faut enfin construire l'état antisymétrique Ψ_A qui, pour deux fermions, se présente encore après antisymétrisation sous la forme factorisée, variables d'espace d'un côté, variables de spin de l'autre (voir (22.83) et (22.86)). Les fonctions de spin sont les X_{S,M_S} :

- état triplet, dont les composantes sont symétriques dans l'échange :

$$X_{1,+1} = \alpha(m_{s_1})\alpha(m_{s_2}) \ , \qquad X_{1,-1} = \beta(m_{s_1})\beta(m_{s_2}) \ , \tag{27.52}$$

[31] E_I désigne toujours l'énergie d'ionisation de l'hydrogène.

[32] De façon imagée : chaque électron a envie de chasser l'autre. Quand l'un est excité, l'autre a juste à *pousser* un peu...

$$X_{1,0} = \frac{1}{\sqrt{2}}\left[\alpha(m_{s_1})\beta(m_{s_2}) + \beta(m_{s_1})\alpha(m_{s_2})\right] \ . \tag{27.53}$$

- état singulet, dont la composante est antisymétrique :

$$X_{0,0} = \frac{1}{\sqrt{2}}\left[\alpha(m_{s_1})\beta(m_{s_2}) - \beta(m_{s_1})\alpha(m_{s_2})\right] \ . \tag{27.54}$$

Pour l'état triplet, la bonne fonction d'espace est :

$$\Phi_1(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}}\left[\phi_n(\vec{r}_1)\phi_m(\vec{r}_2) - \phi_m(\vec{r}_1)\phi_n(\vec{r}_2)\right] \ , \tag{27.55}$$

tandis que la partie spatiale du singulet est :

$$\Phi_0(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}}\left[\phi_n(\vec{r}_1)\phi_m(\vec{r}_2) + \phi_m(\vec{r}_1)\phi_n(\vec{r}_2)\right] \ . \tag{27.56}$$

Les états de spin étant déterminés, ainsi que la forme de la fonction d'onde, il convient de préciser la partie spatiale. L'usage désigne par 1s, 2s, 2p, etc., les états hydrogénoïdes que l'on note ici ϕ_{nlm_l}, pour rappeler qu'il s'agit d'orbitales. Avec deux électrons, les configurations atomiques se forment en choisissant deux états de ce type ; dans l'ordre des énergies croissantes et compte tenu de la dégénérescence accidentelle du champ Coulombien, les configurations iso-énergétiques sont :

$$\begin{array}{llll} (1\text{s})^2 & & & \\ (1\text{s})\,(2\text{s}) & (1\text{s})\,(2\text{p}) & & \\ (1\text{s})\,(3\text{s}) & (1\text{s})\,(3\text{p}) & (1\text{s})\,(3\text{d}) & \text{etc.} \end{array} \tag{27.57}$$

Pour l'hélium, à la configuration $(1\text{s})^2$ correspond le produit $\phi_n\phi_m = \phi_{100}\,\phi_{100}$ et ainsi de suite :

$$\begin{array}{lll} (1\text{s})^2 \ : \phi_{100}\,\phi_{100} & & \\ (1\text{s})\,(2\text{s}) \ : \phi_{100}\,\phi_{200} \ , & (1\text{s})\,(2\text{p}) \ : \phi_{100}\,\phi_{21m_l} & \\ (1\text{s})\,(3\text{s}) \ : \phi_{100}\,\phi_{300} \ , & (1\text{s})\,(3\text{p}) \ : \phi_{100}\,\phi_{31m_l} \ , & (1\text{s})\,(3\text{d}) \ : \phi_{100}\,\phi_{32m_l} \ , \quad \text{etc.} \end{array} \tag{27.58}$$

Les états propres de H_0 sont complètement déterminés une fois la partie d'espace choisie. On remarque que si ϕ_n et ϕ_m sont une seule et même fonction orbitale ϕ_{nlm_l}, seul l'état singulet existe, en conformité avec le Principe de Pauli (techniquement : alors la partie spatiale de l'état triplet est identiquement nulle) ; dit autrement, si $\phi_n \equiv \phi_{nlm_l}$, les trois nombres quantiques orbitaux sont les mêmes pour les deux électrons, qui doivent donc avoir un spin différent, comme on dit ; ceci n'est pas possible pour l'état triplet.

Dans cette description initiale, l'état fondamental est un singulet S ($L = 0$) et est noté 1^1S :

$$\Psi_{1^1\text{S}}(1, 2) = \phi_{100}(\vec{r}_1)\phi_{100}(\vec{r}_2)\,X_{0,0} \ , \tag{27.59}$$

et son énergie vaut $2E_1 \simeq -108,9$ eV. Le premier niveau excité est formé sur la configuration (1s)(2s) ou (1s)(2p), et contient à chaque fois un état triplet et un état singulet, puisque les deux orbitales ϕ_n et ϕ_m sont maintenant différentes :

$$\text{triplet :} \quad \frac{1}{\sqrt{2}}[\phi_{100}\phi_{2lm_l} - \phi_{2lm_l}\phi_{100}]\, X_{1,M_S} \tag{27.60}$$

$$\text{singulet :} \quad \frac{1}{\sqrt{2}}[\phi_{100}\phi_{2lm_l} + \phi_{2lm_l}\phi_{100}]\, X_{0,0} \ , \tag{27.61}$$

avec à chaque fois $l = 0,\, 1,\, -l \leq m_l \leq l$. Ce niveau d'ordre zéro a une énergie égale à $E_1 + E_2 \simeq -54,45 - \frac{54,45}{4} \simeq -68,06$ eV (voir (27.50)) et a une dégénérescence d'ordre 16 : il y a 4 composantes de spin (3 + 1) et 4 possibilités pour le couple (l, m_l). Le trou de Fermi est visible pour l'état triplet : la fonction $\phi_{100}(\vec{r}_1)\phi_{2lm_l}(\vec{r}_2) - \phi_{2lm_l}(\vec{r}_1)\phi_{100}(\vec{r}_2)$ est manifestement nulle si $\vec{r}_1 = \vec{r}_2$.

Si on choisit la configuration (1s)(2s), on a encore un état de type S ($L = 0$) ; d'où un singulet S et un triplet S, notés spectroscopiquement 2^1S et 1^3S, épuisant 4 directions dans le sous-espace dégénéré. De même, en prenant (1s)(2p), on forme deux états de type P ($L = 1$), notés 1^1P et 1^3P dont les dégénérescences sont 3 et 9, et ainsi de suite.

La description précédente, à particules indépendantes, n'est qu'un point de départ. L'inclusion la plus simple de la répulsion entre les électrons consiste à appliquer la théorie des perturbations stationnaires en traitant $V = \frac{e'^2}{r_{12}}$ comme une perturbation. On verra après coup que la répulsion est loin d'être une petite correction (ce qui n'est guère surprenant physiquement) – d'où quelque suspicion sur la validité d'un traitement perturbatif –, mais la prise en compte, même grossière, de la répulsion de Coulomb lève le voile sur un phénomène inattendu, déjà mentionné : en dépit de l'absence d'interactions magnétiques (le Hamiltonien est purement *électrostatique*), les états atomiques ont, grâce au Principe de Pauli et le spin qui va avec, une énergie qui dépend explicitement de la valeur du spin total de l'atome. De toute façon, en pratique et quand il le faut, on peut toujours raffiner le calcul afin d'obtenir une description numériquement plus convenable.

En ce qui concerne l'état fondamental (non-dégénéré), la correction d'énergie au premier ordre est simplement :

$$\Delta E_{1^1\mathrm{S}} = \langle \Psi_{1^1\mathrm{S}} | \frac{e'^2}{r_{12}} | \Psi_{1^1\mathrm{S}} \rangle \ . \tag{27.62}$$

L'opérateur de perturbation ne dépend pas du spin ; après sommation sur les variables de spin, on obtient une intégrale portant sur les seules variables orbitales :

$$\Delta E_{1^1\mathrm{S}} = \langle \phi_{100}\phi_{100} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{100} \rangle \ , \tag{27.63}$$

où l'élément de matrice est l'intégrale :

$$\int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2\, \phi_{100}(r_1)\phi_{100}(r_2)\, \frac{e'^2}{r_{12}}\, \phi_{100}(r_1)\phi_{100}(r_2) \equiv J_{1\mathrm{s}\,1\mathrm{s}} \ . \tag{27.64}$$

L'intégrale $J_{1s\,1s}$ s'appelle intégrale directe[33] et représente la répulsion électrostatique des deux distributions de charges de densités respectives $|\phi_{100}(\vec{r}_1)|^2$ et $|\phi_{100}(\vec{r}_2)|^2$. Montrons en détail comment calculer ce type d'intégrale ; comme on l'a déjà vu pour arriver en (27.44), il suffit de développer $\frac{1}{r_{12}}$ en série de produits d'harmoniques sphériques – ce qui sépare les variables $\vec{r}_1$ et $\vec{r}_2$. Compte tenu de (18.123), une intégrale directe avec ϕ_{nlm} et $\phi_{n'l'm'}$ s'écrit :

$$J_{nl,n'l'} = 4\pi e'^2 \sum_{k\in\mathbb{N}} \sum_{q=-k}^{+k} \frac{1}{2k+1} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \, |R_{nl}(r_1)|^2 |R_{n'l'}(r_2)|^2 \, \frac{r_<^k}{r_>^{k+1}} \times$$
$$Y_{kq}(\theta_1,\phi_1) Y_{kq}^*(\theta_2,\phi_2) Y_{lm}^*(\theta_1,\phi_1) Y_{lm}(\theta_1,\phi_1)\, Y_{l'm'}^*(\theta_2,\phi_2) Y_{l'm'}(\theta_2,\phi_2) \ . \qquad (27.65)$$

Chaque élément de volume est $r_i^2 \mathrm{d}r_i \, \sin\theta_i \mathrm{d}\phi_i$; comme les variables angulaires sont maintenant séparées, on rencontre en fait le produit $\langle Y_{lm}|Y_{kq}|Y_{lm}\rangle\langle Y_{l'm'}|Y_{kq}^*|Y_{l'm'}\rangle$; l'intégration sur les angles ϕ_i donne $\delta_{q0}\delta_{q0}$: seul subsiste le terme $q=0$ de la somme interne, et on est réduit à trouver ce que vaut $\langle Y_{l0}|Y_{k0}|Y_{l0}\rangle\langle Y_{l'0}|Y_{k0}|Y_{l'0}\rangle$. Chaque élément de matrice est nul par parité[34] si k est impair. En outre, chacun d'entre eux peut être vu comme le produit scalaire d'un moment cinétique l (resp. l') avec la somme de deux moments cinétiques k et l (resp. k et l') ; il faut donc en plus que les inégalités triangulaires soient satisfaites, soit $|k-l| \leq l \leq k+l$, ce qui en définitive impose k pair et $k \leq 2l$. Au total, pour une intégrale directe, les seuls termes donnant une contribution non-nulle à la somme sur k sont ceux satisfaisant k pair et $k \leq \min(l, l')$.

Cela étant dit, il reste à préciser l'intégrale radiale, que l'on coupe en deux morceaux pour expliciter le rapport $\frac{r_<^k}{r_>^{k+1}}$; on écrit ainsi :

$$\int_0^{+\infty} r_1^2 \mathrm{d}r_1 \left[\int_0^{r_1} r_2^2 \mathrm{d}r_2 |R_{nl}(r_1)|^2 |R_{n'l'}(r_2)|^2 \frac{r_2^k}{r_1^{k+1}} + \int_{r_1}^{+\infty} r_2^2 \mathrm{d}r_2 |R_{nl}(r_1)|^2 |R_{n'l'}(r_2)|^2 \frac{r_1^k}{r_2^{k+1}} \right] \ . \qquad (27.66)$$

Les variables r_1 et r_2 étant maintenant séparées, ceci s'écrit comme une somme de produits de deux intégrales :

$$\int_0^{+\infty} r_1^{1-k} \mathrm{d}r_1 |R_{nl}(r_1)|^2 \int_0^{r_1} r_2^{k+2} \mathrm{d}r_2 |R_{n'l'}(r_2)|^2 +$$
$$\int_0^{+\infty} r_1^{k+2} \mathrm{d}r_1 |R_{nl}(r_1)|^2 \int_{r_1}^{+\infty} r_2^{1-k} \mathrm{d}r_2 |R_{n'l'}(r_2)|^2 \ . \qquad (27.67)$$

Compte tenu de la Remarque 2 p. 1240, il suffit de calculer ces intégrales I pour *une* valeur du numéro atomique Z, pour être en mesure d'en déduire toutes les autres paramétrées par $Z' \neq Z$: $I(Z') = \frac{Z'}{Z} I(Z)$.

[33] *Attention !* Ceci est la terminologie usuelle en physique atomique et moléculaire (voir par exemple [89]). En revanche, Cohen-Tannoudji et al. [25] notent K l'intégrale directe et J l'intégrale d'échange (traditionnellement, une intégrale d'échange est notée J en physique de la matière condensée).

[34] On rappelle que Y_{lm} a la parité $(-1)^l$.

Dans le cas de $J_{1s\,1s}$, on a précisément :

$$J_{1s\,1s} = e'^2 \int_0^{+\infty} r_1 \mathrm{d}r_1 |R_{10}(r_1)|^2 \int_0^{r_1} r_2^2 \mathrm{d}r_2 |R_{10}(r_2)|^2 + e'^2 \int_0^{+\infty} r_1^2 \mathrm{d}r_1 |R_{10}(r_1)|^2 \int_{r_1}^{+\infty} r_2 \mathrm{d}r_2 |R_{10}(r_2)|^2 \quad . \quad (27.68)$$

Connaissant les fonctions radiales hydrogénoïdes, un calcul simple donne $J_{1s\,1s} = \frac{5}{4}\frac{e'^2}{a_0}$, soit $J_{1s\,1s} \simeq 34,01$ eV ; c'est évidemment une quantité positive... et c'est une correction très importante ! À cet ordre, l'énergie du fondamental est :

$$\boxed{E_{1^1\mathrm{S}} = 2E_1 + J_{1s\,1s} \simeq -74,89\,\mathrm{eV}} \qquad (27.69)$$

La répulsion de Coulomb entre électrons déstabilise considérablement le fondamental, ce qui n'est pas surprenant. La valeur expérimentale étant voisine de $-79,01$ eV, l'erreur est d'environ 5%.

Le premier niveau excité d'énergie E_1+E_2 (à l'ordre zéro) est fortement dégénéré : $2 \times (2 \times 2^2) = 16$ fois ; selon la théorie des perturbations, que l'on a décidé de mettre en œuvre, il convient donc d'abord d'écrire la matrice de la perturbation réduite à ce sous-espace dégénéré, puis de la diagonaliser. En fait, la matrice est déjà diagonale car $V = \frac{e'^2}{r_{12}}$ est invariant par rotation des variables d'espace et ne dépend pas du spin ; il en résulte :

$$[V, \vec{L}] = 0 \ , \qquad [V, \vec{S}] = 0 \ , \qquad (27.70)$$

de sorte que V est un opérateur scalaire vis-à-vis de l'espace et du spin séparément. Pour deux états $\Psi_{LSM_LM_S}$ on a donc nécessairement :

$$\langle \Psi_{LSM_LM_S} | \frac{e'^2}{r_{12}} | \Psi_{L'S'M'_LM'_S} \rangle \propto \delta_{LL'}\,\delta_{SS'}\,\delta_{M_LM'_L}\,\delta_{M_SM'_S} \ , \qquad (27.71)$$

et ceci montre que la matrice de la perturbation est d'emblée diagonale ; les corrections à l'énergie au premier ordre sont simplement données par ces éléments diagonaux, qui vont différer suivant les valeurs de L et S : la répulsion de Coulomb va lever partiellement la dégénérescence présente dans l'approximation grossière des électrons indépendants[35].

Examinons d'abord la configuration (1s)(2s), avec laquelle on ne peut former qu'un état $L = 0$, et deux états de spin $S = 0$ et $S = 1$. En ce qui concerne l'état singulet, noté 2^1S, la correction $\Delta E_{2\,^1\mathrm{S}}$ vaut $\langle \Psi_{2^1\mathrm{S}} | \frac{e'^2}{r_{12}} | \Psi_{2^1\mathrm{S}} \rangle$; après sommation sur les variables de spin, elle prend la forme :

$$\Delta E_{2^1\mathrm{S}} = \frac{1}{2} \langle \phi_{100}\phi_{200} + \phi_{200}\phi_{100} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{200} + \phi_{200}\phi_{100} \rangle \ , \qquad (27.72)$$

soit :

$$\Delta E_{2^1\mathrm{S}} = \langle \phi_{100}\phi_{200} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{200} \rangle + \langle \phi_{100}\phi_{200} | \frac{e'^2}{r_{12}} | \phi_{200}\phi_{100} \rangle \equiv J_{1s\,2s} + K_{1s\,2s} \ . \quad (27.73)$$

[35] Bien sûr, les énergies perturbées ne peuvent dépendre ni de M_L, ni de M_S, puisqu'il n'y a pas de direction privilégiée dans l'espace.

On retrouve une intégrale du type direct, $J_{1s\,2s}$ (répulsion électrostatique entre les deux distributions $|\phi_{100}|^2$ et $|\phi_{200}|^2$, qui est égale à $\frac{34}{81}\frac{e'^2}{a_0}$, soit environ 11,42 eV), mais apparaît maintenant en plus une autre intégrale, $K_{1s\,2s}$, appelée intégrale d'échange[36], égale à $\frac{32}{729}\frac{e'^2}{a_0} \simeq 1,19$ eV.

Cela ne saute pas aux yeux mais il est possible de montrer qu'une intégrale d'échange[37] est toujours positive ([215], Appendice 19, et ci-après problème 27.8.3 pour une autre démonstration) :

$$K_{nm} = \int \mathrm{d}^3r_1 \int \mathrm{d}^3r_2\, \phi_n^*(\vec{r}_1)\phi_m^*(\vec{r}_2)\,\frac{e'^2}{r_{12}}\,\phi_m(\vec{r}_1)\phi_n(\vec{r}_2) > 0 \quad \forall\phi_n,\,\phi_m \ . \qquad (27.74)$$

Noter qu'une telle expression ne peut recevoir d'interprétation classique en termes d'interaction entre deux distributions continues de charge, puisque les produits $\phi_n^*(\vec{r}_1)\phi_m(\vec{r}_1)$ et $\phi_n^*(\vec{r}_2)\phi_m(\vec{r}_2)$ n'ont aucune raison d'être positifs, et ne le sont pas en général.

À cet ordre d'approximation, on obtient :

$$E_{2^1\mathrm{S}} = E_1 + E_2 + J_{1s\,2s} + K_{1s\,2s} = -\frac{2969}{729}\,\frac{e'^2}{2a_0} \simeq -55,41\,\mathrm{eV} \qquad (27.75)$$

qui donne un état à environ 1 eV au-dessous du fondamental de He^+.

Le calcul se fait de la même façon pour l'état triplet 1^3S ; compte tenu du changement de signe dans la partie d'espace des fonctions d'onde entre 2^1S et 1^3S, on trouve maintenant :

$$E_{1^3\mathrm{S}} = E_1 + E_2 + J_{1s\,2s} - K_{1s\,2s} = -\frac{3097}{729}\,\frac{e'^2}{2a_0} \simeq -57,80\,\mathrm{eV}\ , \qquad (27.76)$$

et, compte tenu du signe de l'intégrale d'échange, (27.74), on obtient l'inégalité importante :

$$E_{1^3\mathrm{S}} < E_{2^1\mathrm{S}}\ , \qquad E_{2^1\mathrm{S}} - E_{1^3\mathrm{S}} = 2K_{1s\,2s} \simeq 2,39\,\mathrm{eV} > 0\ , \qquad (27.77)$$

qui est la traduction énergétique du trou de Fermi, prévue qualitativement dans le chapitre 22 (voir p. 1013).

Ainsi apparaît, dans le cadre le plus simple possible, un résultat capital[38] : bien qu'il n'y ait aucune interaction magnétique *explicite*, l'énergie dépend de la valeur du

[36] L'apparition du terme d'échange est clairement une conséquence du postulat de symétrisation et n'a pas d'interprétation classique. On note un bon facteur 10 entre les intégrales directe et d'échange.

[37] L'origine de la terminologie est la suivante : en imaginant que l'on peut préparer l'atome initialement dans l'état $\phi_n(\vec{r}_1)\phi_m(\vec{r}_2)$, alors celui-ci oscille entre cet état et $\phi_m(\vec{r}_1)\phi_n(\vec{r}_2)$ à la fréquence $h^{-1}K_{nm}$. On peut aussi dire que $\hbar^{-1}K_{nm}$ est la fréquence à laquelle *"les électrons échangent leurs nombres quantiques"*. Pour la configuration (1s)(2p), l'échange a lieu toutes les $\sim 10^{-14}$ s ; en revanche, pour la configuration (1s)(n = 10, l = 9), le temps d'échange est égal à ... 1600 ans, alors que le diamètre de l'orbite excitée est seulement de l'ordre de $100\,a_0$ [89]. K provient du postulat de symétrisation : la symétrisation est donc inutile si les fonctions d'onde ne se recouvrent pas [15].

[38] Pour une belle présentation des mille facettes du spin, voir l'article de Frölich [216].

Autre résultat important : sans le spin et le Principe de Pauli, la molécule H_2 n'existerait pas (voir chapitre 28, notamment fig. 28.11)...

spin total : c'est le *magnétisme sans magnétisme* annoncé, formule destinée à marquer l'esprit pour attirer l'attention sur l'origine première du magnétisme. Bien sûr, il y a du *magnétisme* dans la théorie, importé par le spin de l'électron, lequel n'a de sens qu'en tant que conséquence d'un moment magnétique propre révélé par l'expérience. Il n'empêche que le spin à lui seul ne suffit pas : outre l'interaction *électrostatique*, il faut en plus le Principe de Pauli, qui donne une symétrie différente aux fonctions d'espace selon que le spin total est 0 ou 1 et entraîne que l'intégrale d'échange vient avec un signe et le signe contraire, suivant le cas. Toutes choses égales par ailleurs, et compte tenu de la positivité de l'intégrale d'échange, l'énergie du triplet est inférieure à celle du singulet. La dépendance de l'énergie par rapport au spin est un fait majeur, autant sur le plan qualitatif que quantitatif : elle rend compte fondamentalement des propriétés magnétiques de la matière. Des interactions magnétiques "classiques" non seulement ne permettraient pas de comprendre les phénomènes observés mais de plus seraient bien incapables d'expliquer les ordres de grandeur constatés : comme on l'a vu, ces interactions contribuent aux écarts de structure fine et sont, de ce fait dans la gamme d'énergie $\sim 10^{-4}$ eV, alors que l'intégrale d'échange est de l'ordre de l'eV.

Ce point important ayant été mis en avant, on voit, en résumé, que pour la configuration (1s)(2s), l'approximation en cours donne deux états excités d'énergies :

$$\boxed{E_{1^3\mathrm{S}} = E_1 + E_2 + J_{1\mathrm{s}\,2\mathrm{s}} - K_{1\mathrm{s}\,2\mathrm{s}} \simeq -57,80\,\mathrm{eV}} \tag{27.78}$$

$$\boxed{E_{2^1\mathrm{S}} = E_{1^3\mathrm{S}} + 2K_{1\mathrm{s}\,2\mathrm{s}} \simeq -55,41\,\mathrm{eV}} \tag{27.79}$$

Pour les états P on trouve de façon analogue[39] :

$$E_{1^1\mathrm{P}} = E_1 + E_2 + J_{1\mathrm{s}\,2\mathrm{p}} + K_{1\mathrm{s}\,2\mathrm{p}} = -\frac{25859}{6561}\,\frac{e'^2}{2a_0} \simeq -53,62\,\mathrm{eV} \ , \tag{27.80}$$

$$E_{1^3\mathrm{P}} = E_1 + E_2 + J_{1\mathrm{s}\,2\mathrm{p}} - K_{1\mathrm{s}\,2\mathrm{p}} = -\frac{26755}{6561}\,\frac{e'^2}{2a_0} \simeq -55,48\,\mathrm{eV} \ , \tag{27.81}$$

avec à nouveau :

$$E_{1^3\mathrm{P}} < E_{1^1\mathrm{P}} \ , \qquad E_{1^1\mathrm{P}} - E_{1^3\mathrm{P}} = 2K_{1\mathrm{s}\,2\mathrm{p}} > 0 \ . \tag{27.82}$$

En résumé, pour la configuration (1s)(2p), on a :

$$\boxed{E_{1^3\mathrm{P}} = E_1 + E_2 + J_{1\mathrm{s}\,2\mathrm{p}} - K_{1\mathrm{s}\,2\mathrm{p}} \simeq -55,48\,\mathrm{eV}} \tag{27.83}$$

[39] Pour l'intégrale directe, c'est à nouveau le terme $k = 0$ qui compte, avec toujours $q = 0$ (et le facteur angulaire vaut $\frac{1}{4\pi}$, qui se compense avec $\frac{4\pi}{2\times 0+1}$) ; on trouve ainsi $J_{1\mathrm{s}\,2\mathrm{p}} = \frac{361}{729}\frac{e'^2}{a_0} \simeq 3,48$ eV. Pour l'intégrale d'échange, c'est le terme $k = 1$ qui compte ; le facteur global est $\frac{4\pi}{2\times 1+1}\frac{1}{4\pi} = \frac{1}{3}$; on trouve ainsi $K_{1\mathrm{s}\,2\mathrm{p}} = \frac{224}{6561}\frac{e'^2}{a_0} \simeq 0,92$ eV.

En règle générale, si l'intégrale d'échange K est plus petite que l'intégrale directe associée J c'est, essentiellement, parce que dans cette dernière l'intégrand est toujours positif, alors qu'il peut être négatif pour K.

$$\boxed{E_{1^1\mathrm{P}} = E_{1^3\mathrm{P}} + 2K_{1\mathrm{s}\,2\mathrm{p}} \simeq -53,62\,\mathrm{eV}} \tag{27.84}$$

En définitive, la succession des niveaux classés par énergie croissante est, dans l'approximation en cours (énergies en eV) :

$$E_{1^1\mathrm{S}} \simeq -74,9 < E_{1^3\mathrm{S}} \simeq -57,8 < E_{1^3\mathrm{P}} \simeq -55,5 < E_{2^1\mathrm{S}} \simeq -55,4 < E_{1^1\mathrm{P}} \simeq -53,6 \ . \tag{27.85}$$

Au total, l'approximation assez grossière[40] donne des résultats moyennement satisfaisants. Si la dégénerescence d'échange est bien levée, l'ordre des niveaux n'est pas conforme à l'expérience[41] ; du spectre observé, on peut en effet conclure que l'ordre réel est le suivant :

$$E_{1^1\mathrm{S}} \simeq -79,0 < E_{1^3\mathrm{S}} \simeq -59,2 < E_{2^1\mathrm{S}} \simeq -58,4 < E_{1^3\mathrm{P}} \simeq -58,1 < E_{1^1\mathrm{P}} \simeq -57,8 \ . \tag{27.86}$$

Les niveaux 2^1S et 1^3P sont très proches l'un de l'autre ($\lesssim 0,3$ eV, à peine 1% en valeur relative), et sont mal rendus par une méthode trop rustique dont l'erreur typique est de l'ordre de 5%.

Quelle que soit la description précise utilisée, un fait est sûr : avec ses deux électrons, l'atome d'hélium possède deux familles d'états, ceux où le spin total est nul et ceux où il vaut 1. On sait que les transitions les plus intenses sont du type dipolaire électrique (E1), et que l'une des règles de sélection correspondantes est $\Delta S = 0$. Ainsi, quand on n'est pas en mesure d'observer les transitions dipolaires magnétiques – parce qu'elles sont de faible intensité –, tout se passe comme si les deux ensembles d'états s'ignoraient l'un l'autre. C'est pourquoi, historiquement, on en est venu à distinguer *deux* sortes d'hélium[42] : la variété *ortho* (spin $S = 1$) et la variété *para* ($S = 0$), avec des spectres fort différents... et des propriétés magnétiques contrastées (seul l'orthohélium est paramagnétique) ; en outre, le "fondamental" de l'orthohélium a une durée de vie très longue, le retour au *fondamental* de l'hélium se faisant par une transition très rare ; cette durée de vie est difficile à mesurer en raison de la sensibilité de cet état aux collisions, et plus généralement aux conditions physiques de la vapeur.

La prise en compte de la structure fine peut se tenter suivant la méthode décrite dans la section 27.2, voir notamment (27.24). Dans ce cadre, les états ^{1}S ne bougent pas, les ^{3}S sont déplacés légèrement, mais il n'y a pas de structure fine à proprement parler puisque, pour chaque terme spectral, J ne peut prendre qu'une seule valeur, tout comme pour ^{1}P d'ailleurs. En revanche, le terme spectral ^{3}P donne les trois valeurs 0, 1 et 2 pour J et éclate donc en trois niveaux de structure fine. Numériquement, l'accord n'est pas très bon, à cause de l'importance de l'interaction dipolaire magnétique spin - spin (négligée dans le schéma avec $A\vec{L}.\vec{S}$) ; sa prise en compte donne des résultats acceptables, mais loin d'être pleinement satisfaisants en raison de la précision des mesures spectroscopiques, qui exigent la prise en compte théorique de toutes les corrections relativistes.

---

[40] On traite par perturbation et au premier ordre les effets d'un opérateur dont les valeurs moyennes ne sont pas très petites vis-à-vis des écarts de niveaux du problème non perturbé.

[41] On peut trouver un schéma des niveaux de l'hélium à l'adresse suivante :
`http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/helium.`

[42] Moyen anti-mnémotechnique : *para* comme non paramagnétique...

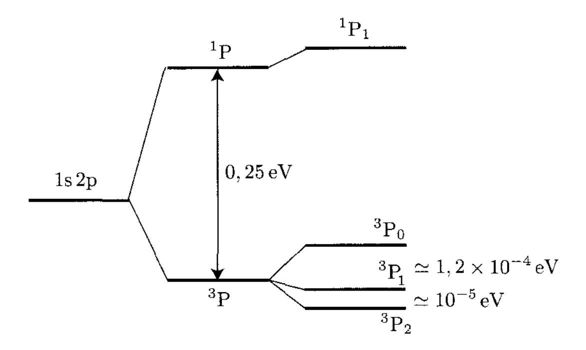


Figure 27.1: Structure fine de He pour les multiplets P issus de (1s)(2p) (pour plus de détails, voir [25]).

Expérimentalement, les transitions $^3S_1 \rightleftarrows {}^3P_{J=0,1,2}$ ont été très soigneusement étudiées ; on observe les valeurs de la table 27.1, d'où l'on déduit les écarts de structure fine : $E(^3P_0) - E(^3P_1) \simeq 1,247 \times 10^{-4}$ eV et $E(^3P_1) - E(^3P_0) \simeq 9,511 \times 10^{-6}$ eV. L'étude détaillée de la structure fine du multiplet ^3P présente un grand intérêt sur le plan fondamental, car elle est un cas d'école jouant un rôle majeur pour la détermination précise de la constante de structure fine α.

Table 27.1: La raie He à 1083 nm

transition	longueur d'onde (nm)
$^3S_1 \rightleftarrows {}^3P_0$	1082,907
$^3S_1 \rightleftarrows {}^3P_1$	1083,025
$^3S_1 \rightleftarrows {}^3P_1$	1083,034

27.8 Exercices et problèmes

27.8.1 Opérateur effectif spin - spin pour deux électrons

Pour un système à $N = 2$ électrons, le choix de deux orbitales distinctes ϕ_n et $\phi_{n'}$ permet toujours de construire un état singulet et un état triplet.

1. En utilisant la théorie des perturbations au premier ordre pour traiter la répulsion entre électrons, exprimer les énergies des états triplets et singulet à l'aide des intégrales directe J et d'échange K.

2. En déduire que, restant dans le sous-espace d'états construit avec les deux seules fonctions ϕ_n et $\phi_{n'}$, il est équivalent d'introduire un opérateur effectif du genre $V_{\text{eff}} = -A\vec{S}_1.\vec{S}_2$, où A est une constante positive, à exprimer en fonction de K.

Cette équivalence est à la base des Hamiltoniens-modèles souvent rencontrés en Physique de la matière condensée (modèle de Heisenberg, par exemple, voir chapitre 30, sous-section 30.6.1, éq. 30.202).

♦ *Remarque*

Bien évidemment, une telle formulation ne doit pas laisser croire qu'il s'agit d'interactions magnétiques au sens classique du terme : à nouveau, l'ordre de grandeur énergétique n'y serait pas. Il s'agit tout juste d'une représentation effective, en terme d'opérateurs, du *splitting* singulet - triplet obtenu plus haut. Dans le même ordre d'idée, rien n'interdit de représenter un atome à deux niveaux dans un langage de *spin fictif* $S = \frac{1}{2}$. ♦

27.8.2 L'ion H^- existe-t-il ?

Avec les éléments donnés dans ce chapitre, discuter la stabilité de l'ion[43] H^-.

Que pensez-vous de ces résultats sachant que cet ion n'a qu'un seul état lié [217], d'énergie $\simeq -0,75$ eV ([89], [218]) ?

27.8.3 Positivité d'une intégrale d'échange

Une intégrale d'échange est par définition de la forme :

$$K \stackrel{\text{déf}}{=} \int \mathrm{d}^3r_1\mathrm{d}^3r_2\,\phi_a^*(\vec{r}_1)\phi_b^*(\vec{r}_2)\frac{1}{r_{12}}\phi_b(\vec{r}_1)\phi_a(\vec{r}_2) \equiv \int \mathrm{d}^3r_1\mathrm{d}^3r_2\,f_{ab}^*(\vec{r}_1)\frac{1}{r_{12}}f_{ab}(\vec{r}_2)\ . \tag{27.87}$$

1. La transformée de Fourier d'une fonction $f(\vec{r})$ est $F(\vec{k}) \stackrel{\text{déf}}{=} \int \mathrm{d}^3r\,f(\vec{r})\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$, la formule inverse étant $f(\vec{r}) = (2\pi)^{-3}\int \mathrm{d}^3k\,F(\vec{k})\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}$. Trouver la transformée de Fourier $C(\vec{k})$ de $\frac{1}{r}$, prise comme la limite $\lambda \to 0_+$ du potentiel de Yukawa $\frac{1}{r}\mathrm{e}^{-\lambda r}$.

2. À l'aide du théorème de convolution, et en utilisant la formule inverse de Fourier, montrer que :

$$\int \mathrm{d}^3r_2\,f_{ab}(\vec{r}_2)\frac{1}{r_{12}} = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}_1}\,F_{ab}(\vec{k})C(\vec{k})\ , \tag{27.88}$$

où $F_{ab}(\vec{k})$ est la transformée de Fourier de $f_{ab}(\vec{r})$.

3. En déduire que K est égale à l'intégrale sur $\vec{k}$ d'une fonction positive, et conclure.

[43] Cet ion est présent dans les atmosphères d'étoile, dont le Soleil.

27.8.4 États de moments angulaires donnés associés à une configuration électronique

Un atome à $Z+2$ électrons a pour configuration électronique $(\mathrm{CC})(nl)^x(n'l')^{x'}$ où (CC) désigne l'ensemble des couches complètes occupées par Z électrons. On désigne par $\vec{L}$, $\vec{S}$ et $\vec{J}$ les moments cinétiques totaux.

1. Quelles sont les valeurs possibles de Z de cet atome hypothétique ?
2. Trouver N_{D}, nombre de déterminants distincts que l'on peut construire à partir d'une telle configuration.
3. Quelle est la dégénérescence de rotation, N_{LS}, liée aux moments $\vec{L}$ et $\vec{S}$?
4. Quelle conclusion peut-on tirer si $N_{\mathrm{D}} < N_{LS}$?
5. L et S étant donnés, quelles sont les valeurs possibles de J ?
6. Soit N_J la dégénérescence totale de rotation pour l'ensemble des valeurs possibles de J. Quelle est la relation entre N_J et N_{LS} ?
7. On choisit $x_1 = 2$, $x_2 = 0$.
   (a) Combien valent N_{LS} et N_J ?
   (b) Quelle est la valeur maximale $L_{\max}$ de L ?
   (c) Avec cette valeur, peut-on former un état triplet ?
   (d) Sans faire explicitement les calculs, expliquer comment on peut former les différentes composantes $|L_{\max} S M_L M_S\rangle$.

27.8.5 Structure fine du carbone. Effets Zeeman et Paschen - Back

Le carbone a pour configuration $(1\mathrm{s})^2(2\mathrm{s})^2(2\mathrm{p})^2$.

1. Quelles sont les valeurs possibles de L et S ? Donner la notation correspondante des termes spectraux.
2. Donner un argument qualitatif permettant de penser que l'état de plus basse énergie[44] est $^3\mathrm{P}$.

[44]Ce résultat est en accord avec l'une des règles de Hund affirmant que

> *"Pour une configuration électronique donnée, le terme de plus basse énergie est celui ayant le spin total maximum".*

Une deuxième règle est

> *"Pour un spin total donné, le terme de plus faible énergie est celui ayant la plus grande valeur de L".*

3. Existe-t-il des transitions dipolaires électriques entre ces termes spectraux ?
4. Former les fonctions propres du fondamental.
5. Préciser la structure fine de l'état fondamental, sachant que la constante $A(2,1,1)$ est positive.
6. En présence d'un champ magnétique $\mathcal{B}$ parallèle à Oz, examiner les deux situations relatives aux effets Zeeman et Paschen - Back.

27.8.6 Étude de l'atome d'azote

L'azote a pour configuration électronique fondamentale $(1s)^2(2s)^2(2p)^3$.

1. Quel est le nombre N_{D} de déterminants associés à cette configuration.
2. Quelles sont les valeurs possibles des moments cinétiques orbital et de spin de l'atome ?
3. En procédant par élimination, et par essai et erreur, préciser les termes spectraux associés à cette configuration. Les ordonner en énergie conformément à la règle de Hund.
4. Quelles ont les transitions dipolaires électriques possibles entre ces termes ?
5. On introduit maintenant la structure fine sous la forme (27.24), la constante A étant à chaque fois positive. Préciser les multiplets associés à chaque terme spectral.

27.8.7 Interaction de configurations pour l'atome d'hélium

Pour représenter l'état fondamental de l'atome d'hélium, on forme des combinaisons linéaires de produits de fonctions hydrogénoïdes ($Z=2$), $R_{nl}(r)Y_{lm}(\theta,\phi)$:

$$\Psi(\vec{r}_1,\vec{r}_2) = \sum_{nn'll'mm'} c_{nn'll'mm'}\frac{1}{\sqrt{2}}(\mathbf{1}+T_{12})R_{nl}(r_1)Y_{lm}(\theta_1,\phi_1)R_{n'l'}(r_2)Y_{l'm'}(\theta_2,\phi_2) \ . \tag{27.89}$$

1. Expliquer pourquoi il faut (et il suffit de) prendre $l=l'$.
2. Comment faut-il choisir les nombres m et m' ?
3. En écrivant que la fonction doit être propre de $\vec{L}^2$ avec $L=0$, trouver une relation de récurrence entre les coefficients $A_m \stackrel{\text{déf}}{=} c_{nn'llmm}$.

4. En déduire que la forme générale de la fonction d'onde fondamentale est :

$$\Psi(\vec{r}_1, \vec{r}_2) = \sum_{l \in \mathbb{N}} a_l \, P_l(\cos\theta_{12}) \sum_{n,n'} \frac{1}{\sqrt{2}} (\mathbf{1} + T_{12}) R_{nl}(r_1) R_{n'l}(r_2) \ , \tag{27.90}$$

   où P_l est un polynôme de Legendre et θ_{12} l'angle entre $\vec{r}_1$ et $\vec{r}_2$.

5. Appliquer la méthode précédente avec les deux configurations $(1s)^2$ et $(1s)(2s)$. Comparer aux valeurs expérimentales (voir (27.86)).

6. Expliquer pourquoi une méthode variationnelle, quelle qu'elle soit, pourrait s'appliquer au premier état excité triplet ?

27.8.8 Structure hyperfine du fondamental de l'ion Be$^+$

Analyser et préciser la structure hyperfine du fondamental de l'ion Be$^+$ utilisé pour la mise en évidence de l'effet Zénon (voir chapitre 21, section 21.1, fig. 21.2).

27.8.9 Nature de la transition de la raie 21 cm

Expliquer pourquoi la raie 21 cm est de nature dipolaire magnétique (M1).

physique

Chapitre 28

Introduction à la physique des molécules

Ce chapitre fait le tour d'horizon des notions fondamentales permettant de décrire la physique des molécules. Après la présentation de l'approximation de Born et Oppenheimer, on jette les bases de la théorie de la liaison chimique avant d'en discuter la nature physique. Le chapitre se termine par les rudiments de spectroscopie infrarouge et micro-onde.

28.1 L'approximation de Born et Oppenheimer

Il s'agit ici de présenter l'approximation fondamentale sur laquelle repose la physique moléculaire et la physique de la matière condensée.

28.1.1 Ordres de grandeur

Une molécule (ou un solide) est un assemblage complexe d'électrons (particules peu massives, donc légères) et de noyaux (particules lourdes) en interaction mutuelle, édifice dont on ne sait évidemment pas trouver les états propres exacts. Le bien-fondé de l'approximation fondamentale décrite ci-dessous réside dans la grande différence d'inertie entre électrons et noyaux[1] et conduit à l'approximation dite de Born et Oppenheimer.

[1] Il est souvent physiquement bien plus utile de faire la distinction, non entre les électrons et les noyaux, mais entre *certains* électrons et ce que l'on appelle les *cœurs* ; ces derniers sont une sorte de souvenir tenace des constituants séparés (atomes, ions, ...) et sont eux-même un assemblage robuste d'électrons et de noyaux, peu sensibles en première approximation aux processus physico-chimiques ordinaires. On continuera provisoirement à parler d'électrons et de noyaux, pour la clarté.

Le point central est en effet la grande différence entre la masse électronique m et une masse nucléaire typique, M :

$$\frac{m}{M} \ll 1 \ . \tag{28.1}$$

D'un autre côté, toutes ces particules sont en interaction mutuelle par des forces qui sont toutes du même ordre de grandeur $\sim \frac{e'^2}{a_0^2}$. Des forces comparables mais des inerties très différentes donnent classiquement des mouvements caractérisés par des échelles de temps dont les unes sont courtes et les autres longues : les particules massives se déplacent lentement, les particules légères vont vite. Intuitivement (et de façon imagée), on peut donc affirmer que les électrons ont un mouvement rapide autour des centres attracteurs que constituent les noyaux, lesquels se déplacent lentement.

Corrélativement, l'étude des ordres de grandeur de l'énergie d'une molécule stable permet de dégager simplement trois échelles typiques qui, par le jeu de $E = h\nu$, redonnent tout naturellement plusieurs échelles de temps $\sim \nu^{-1}$ très différentes. Une molécule possède deux sortes de degrés de liberté : les coordonnées électroniques $\{q_i\}$ et les coordonnées des noyaux $\{Q_j\}$. Tous les mouvements sont confinés et se produisent à l'intérieur d'un domaine de dimension atomique, de l'ordre de $2a_0 \simeq 1$ Å. Avec la seule masse de l'électron, l'énergie la plus simple que l'on peut former est $\frac{\hbar^2}{m(2a_0)^2}$ et c'est forcément l'ordre de grandeur de l'énergie associée au mouvement des électrons, E_{el} :

$$\boxed{E_{\text{el}} \sim \frac{\hbar^2}{4ma_0^2} = -\frac{1}{2}\,E_{n=1} \sim 7 \ \text{eV}} \tag{28.2}$$

En ce qui concerne les noyaux, deux types de mouvement sont possibles : les mouvements de vibration et les mouvements de rotation en bloc[2]. Les premiers décrivent les (petites) vibrations effectuées par les noyaux autour de leurs positions d'équilibre, lesquelles définissent la *conformation* de la molécule, au sens de la Chimie (longueurs de liaison, angles de liaison). Ces vibrations sont supposées *petites*, tout simplement parce que si elles prennent de l'amplitude, la molécule se dissocie – or on entend décrire une molécule dûment constituée. Décrites en première approximation comme des oscillateurs harmoniques, elles sont évidemment sujettes à quantification. Enfin, en phase diluée (vapeur) une molécule peut aussi tourner sur elle-même : les angles qui caractérisent son orientation sont des variables cycliques figurant comme arguments dans la fonction d'onde. La nécessité pour celle-ci d'être monovaluée donne lieu à des conditions aux limites qui, par le scénario habituel de la Mécanique quantique, engendre la *quantification* de la rotation moléculaire ; l'exemple le plus simple est le rotateur rigide, constitué d'une molécule diatomique (*haltère*) dont les deux atomes sont à une distance *fixe* l'un de l'autre : après réduction du problème à deux corps, c'est le mouvement quantifié d'une particule ayant la masse réduite et astreinte à se déplacer sur une sphère de rayon donné.

Pour estimer l'énergie liée à la vibration, considérons une molécule diatomique *stable* AB, dont les noyaux sont séparés par la distance R. Physiquement, il doit être possible

[2]On suppose évacué d'emblée le mouvement de translation d'ensemble de la molécule (translation uniforme par rapport à un repère galiléen), non quantifié et sans intérêt.

de définir une certaine fonction $E(R)$ représentant la variation d'énergie lorsque, toutes choses égales par ailleurs, on fait varier la longueur de la liaison AB. Elle a forcément l'allure donnée sur la figure 28.1 : fortement répulsive à courte distance (chacun des partenaires A et B résiste à l'invasion par l'autre de son domaine réservé), présentant un minimum en une certaine valeur R_0 (la molécule est réputée stable) et tendant vers une constante quand R tend vers l'infini – d'où l'identification sans ambiguïté d'une énergie de dissociation, E_{diss}. Expérimentalement, cette énergie est typiquement de l'ordre[3] de quelques eV, soit $E_{\text{diss}} \sim E_{\text{el}}$.

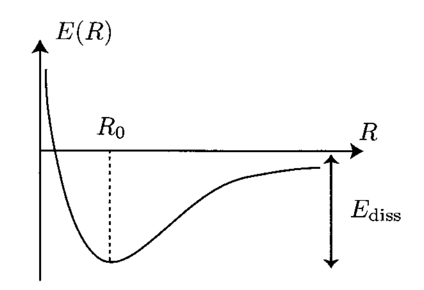


Figure 28.1: Variation typique de la fonction $E(R)$.

S'agissant toujours de dégager des ordres de grandeur, on fait l'hypothèse de petites amplitudes autour de R_0, ce qui permet de confondre le creux de potentiel avec sa parabole osculatrice (voir un exemple sur la figure 28.20), en écrivant :

$$E(R) \simeq E(R_0) + \frac{1}{2}k(R-R_0)^2 \stackrel{\text{déf}}{=} E_{\text{harm}}(R), \; k \stackrel{\text{déf}}{=} \left(\frac{\mathrm{d}^2E}{\mathrm{d}R^2}\right)_{R=R_0}, \; \frac{|R-R_0|}{R_0} \ll 1 \; . \quad (28.3)$$

Quand $R = 2R_0$, on a étiré la liaison de deux fois sa longueur et on peut raisonnablement admettre que l'augmentation d'énergie est du même ordre que l'énergie de dissociation :

$$E(2R_0) \sim E(R_0) + E_{\text{diss}} \quad \Longleftrightarrow \quad \frac{1}{2}\,kR_0^2 \sim \frac{\hbar^2}{4ma_0^2} \; , \qquad (28.4)$$

et comme $R_0 \sim 2a_0$, la constante de raideur est à peu près[4] :

$$k \sim \frac{\hbar^2}{8ma_0^4} \; . \qquad (28.5)$$

Par ailleurs, k est de la forme $M\omega_{\text{vibr}}^2$ où M est une masse nucléaire (par exemple, la masse réduite de deux noyaux), d'où :

$$\omega_{\text{vibr}} \sim \frac{\hbar}{(2a_0)^2\sqrt{mM}} \; . \qquad (28.6)$$

[3]Pour la molécule d'hydrogène, l'énergie de dissociation vaut 4, 71 eV.

[4]Plus directement, on peut aussi écrire que $k \stackrel{\text{déf}}{=} E''(R_0)$ est en ordre de grandeur le rapport d'une énergie électronique et du carré d'une longueur atomique, soit effectivement environ $\frac{\hbar^2}{ma_0^4}$.

Finalement, l'énergie de vibration, $E_{\text{vibr}} = \hbar\omega_{\text{vibr}}$ est environ :

$$\boxed{E_{\text{vibr}} \sim \frac{\hbar^2}{(2a_0)^2\sqrt{mM}}} \tag{28.7}$$

Par comparaison avec (28.2), on voit que :

$$\frac{E_{\text{vibr}}}{E_{\text{el}}} \sim \sqrt{\frac{m}{M}} \ll 1 \ ; \tag{28.8}$$

il convient de noter au passage que si le rapport $\frac{m}{M}$ est en effet très petit, c'est sa racine carrée qui intervient ; elle peut ne pas être si petite que cela ($\sqrt{5\times 10^{-4}} \simeq 2\times 10^{-2}$).

Enfin, trouvons l'ordre de grandeur de l'énergie de rotation, E_{rot}. Elle est du genre $\frac{J^2}{2I}$, où J est un moment cinétique et où I est un moment d'inertie typique, soit $I \sim M(2a_0)^2$. $\hbar$ est l'ordre de grandeur fondamental pour un moment cinétique en Mécanique quantique , donc $J \sim \hbar$, de sorte que :

$$\boxed{E_{\text{rot}} \sim \frac{\hbar^2}{M(2a_0)^2}} \tag{28.9}$$

Par comparaison avec (28.7), il vient :

$$\frac{E_{\text{rot}}}{E_{\text{vibr}}} \sim \sqrt{\frac{m}{M}} \ll 1 \ . \tag{28.10}$$

Ainsi, à chaque fois, c'est le (petit) paramètre de la racine carrée du rapport des masses qui détermine tout. En conclusion de cette analyse en ordres de grandeur :

$$\boxed{\varepsilon \overset{\text{déf}}{=} \sqrt{\frac{m}{M}} \ll 1 \ : \ E_{\text{rot}} \sim \varepsilon E_{\text{vibr}} \sim \varepsilon^2 E_{\text{el}} \quad \Longrightarrow \quad E_{\text{rot}} \ll E_{\text{vibr}} \ll E_{\text{el}}} \tag{28.11}$$

En associant sa pulsation à chaque énergie, on voit que la rotation est lente comparée à la vibration, laquelle est également lente par rapport à la circulation des électrons. En terme de transitions spectrales, si on accepte que les transitions électroniques sont dans l'UV ($\lambda \sim 1000$ Å) et en prenant $m/M \sim 10^{-4}$, on voit que la longueur d'onde d'une transition vibrationnelle est $\lambda_{\text{vibr}} \sim 10^{-9}$ m, soit 10 μ, donnant un spectre dans l'infrarouge. De la même façon, une longueur d'onde rotationnelle sera $\lambda_{\text{rot}} = 1000\ \mu = 1$ mm. En termes de fréquences ($\omega = c/\lambda = 2\pi\nu$) :

$$1000\ \text{Å} \longleftrightarrow \omega_{\text{el}} = 3\times 10^{15}\ \text{rd/s} \quad \Longleftrightarrow \quad \nu_{\text{el}} \simeq 5\times 10^{14}\ \text{Hz} \longleftrightarrow 2\ \text{eV}\ , \tag{28.12}$$

et :

$$\omega_{\text{vibr}} = 3\times 10^{13}\ \text{rd/s} \quad \Longleftrightarrow \quad \nu_{\text{vibr}} \simeq 5\times 10^{12}\ \text{Hz} \longleftrightarrow 0,02\ \text{eV}\ , \tag{28.13}$$

$$\omega_{\text{rot}} = 3\times 10^{11}\ \text{rd/s} \quad \Longleftrightarrow \quad \nu_{\text{rot}} \simeq 5\times 10^{10}\ \text{Hz} = 50\ \text{Ghz} \longleftrightarrow 0,2\ \text{meV}\ . \tag{28.14}$$

Ceci montre notamment que, pour les niveaux de rotation, on ne peut à l'ambiante omettre les facteurs de Boltzmann.

28.1.2 Approximation de Born et Oppenheimer

Compte tenu de l'analyse précédente des ordres de grandeurs, la différence de rapidité des mouvements des électrons et des noyaux suggère de traiter dans une première étape le seul mouvement des électrons, les noyaux étant pris carrément immobiles, quitte à introduire ultérieurement le mouvement lent de ces derniers.

Dans l'approximation électrostatique[5], le Hamiltonien d'une molécule s'écrit :

$$\boxed{H = \sum_j \frac{\vec{P}_j^2}{2M_j} + \sum_i \frac{\vec{p}_i^2}{2m} + V(\{q_i\},\,\{Q_j\}) \equiv T_\text{N} + H_0(p_i,\,q_i\,;\,Q_j)} \tag{28.15}$$

V contient toutes les interactions de Coulomb (électrons - électrons, électrons - noyaux, noyaux - noyaux). La première étape consiste à radicaliser la situation en posant formellement $M_j = +\infty$ dans (28.15), ce qui revient à dire que les noyaux sont fixes. Notons que H_0 ne contient pas d'opérateurs différentiels par rapport aux Q_j, d'où :

$$[H_0\,, Q_j] = 0 \quad \forall\, j\ . \tag{28.16}$$

Étape 1 : mouvement des électrons dans le champ des noyaux fixes

Fixer les noyaux, c'est finalement mettre de côté le terme T_N, et en être réduit à ne devoir chercher que les états propres de H_0. Dans cet opérateur, les p_i et les q_i sont des opérateurs au sens usuel de la Mécanique quantique ; au contraire, les coordonnées nucléaires Q_i sont de simples paramètres. L'équation aux valeurs propres de H_0 en représentation-q s'écrit précisément :

$$H_0\,\Phi(q_i\,;\,Q_j) = E(Q_j)\,\Phi(q_i\,;\,Q_j)\ . \tag{28.17}$$

où $q_i \equiv \{q_i\}_i$, $Q_j \equiv \{Q_j\}_j$. Les coordonnées nucléaires Q_j paramétrisent tout autant les fonctions propres que les valeurs propres. H_0 contient notamment le potentiel attractif des noyaux pour les électrons et, tout comme pour un atome, l'application de conditions aux limites (normalisabilité des fonctions propres des états liés) produit la quantification spontanée des énergies propres. Chaque mode propre est donc repérable par un jeu de nombres quantiques, collectivement notés n. L'équation (28.17) peut donc être écrite plus précisément :

$$H_0\,\Phi_n(q_i\,;\,Q_j) = E_n(Q_j)\,\Phi_n(q_i\,;\,Q_j)\ . \tag{28.18}$$

Résoudre effectivement ce type d'équation est évidemment toute une histoire – admettons cette étape franchie, formellement. En tout état de cause, les Φ_n sont orthogonales, et supposées normalisées à l'unité :

$$\int \prod_i \mathrm{d}q_i\,\Phi_n^*(q_i\,;\,Q_j)\,\Phi_{n'}(q_i\,;\,Q_j) = \delta_{n\,n'}\ ,\forall\, Q_j\ , \tag{28.19}$$

[5] c'est-à-dire sans interactions magnétiques du genre spin-orbite. Bien sûr, pour une molécule (ou un solide) comme pour un atome, les corrections correspondantes doivent être prises en compte dans les traitements qui se veulent plus précis.

où les Q_j sont les *mêmes*[6] dans Φ_n et $\Phi_{n'}$.

Étape 2 : inclusion du mouvement des noyaux

En admettant que les fonctions $\{\Phi_n\}$ ainsi définies constituent une base complète, on peut *a priori* développer la fonction d'onde totale $\Psi(q_i, Q_j)$ représentant la molécule dans sa globalité sous la forme :

$$\Psi(q_i, Q_j) = \sum_n \chi_n(Q_j)\,\Phi_n(q_i\,;\,Q_j)\ . \tag{28.20}$$

Les $\chi_n(Q_j)$, qui constituent les coefficients du développement, ne dépendent par nature que des coordonnées nucléaires. Le report d'une telle expression dans l'équation aux fonctions propres pour la molécule :

$$H\Psi \equiv (T_{\rm N} + H_0)\Psi = E\Psi \tag{28.21}$$

fait apparaître différents termes.

Dans le membre de gauche de (28.21), le premier terme (énergie cinétique des noyaux) contient des dérivées secondes ($\vec{P} = -\mathrm{i}\hbar\vec{\nabla}_Q$) ; en utilisant la formule pour la dérivée seconde d'un produit, $(fg)'' = f''g + 2f'g' + fg''$, il devient :

$$T_{\rm N}\Big(\chi_n(Q_j)\Phi_n(q_i\,;\,Q_j)\Big) = \sum_k \frac{1}{2M_k}\Big[(\vec{P}_k^2\chi_n)\,\Phi_n + 2(\vec{P}_k\chi_n)\,(\vec{P}_k\Phi_n) + \chi_n\,\vec{P}_k^2\Phi_n\Big] =$$

$$(T_{\rm N}\chi_n)\,\Phi_n + \sum_k \frac{1}{M_k}(\vec{P}_k\chi_n)\,(\vec{P}_k\Phi_n) + \chi_n\,(T_{\rm N}\Phi_n)\ . \tag{28.22}$$

Le second terme au premier membre de (28.21) contient les termes du genre :

$$H_0\,\chi_n(Q_j)\Phi_n(q_i\,;\,Q_j) = \chi_n(Q_j)\,H_0\Phi_n(q_i\,;\,Q_j) = E_n(Q_j)\chi_n(Q_j)\Phi_n(q_i\,;\,Q_j)\ , \tag{28.23}$$

où a été utilisé le fait que les Q_j dans H_0 n'interviennent que dans des opérateurs multiplicatifs (pas d'opérateurs cinétiques, voir (28.16)) et que, de ce fait $H_0\chi_n = \chi_n H_0$. Compte tenu de tout ceci et après quelques réarrangements, (28.21) devient :

$$\sum_n \big[(T_{\rm N}\chi_n)\Phi_n + E_n(Q_j)\chi_n\Phi_n\big] = \sum_n \Big[E\chi_n\Phi_n - \sum_k \frac{1}{M_k}(\vec{P}_k\Phi_n)(\vec{P}_k\chi_n) - (T_{\rm N}\Phi_n)\chi_n\Big]\ ; \tag{28.24}$$

multipliant maintenant par Φ_m^*, intégrant sur tous les q_i et tenant compte de (28.19), il vient :

$$T_{\rm N}\chi_m(Q_j) + E_m(Q_j)\,\chi_m(Q_j) = E\,\chi_m(Q_j) - \sum_n \Lambda_{mn}\,\chi_n(Q_j)\ , \tag{28.25}$$

[6] Sinon on aurait deux Hamiltoniens H_0 distincts (pour deux configurations fixes des noyaux), dont les fonctions propres n'ont aucune raison d'être orthogonales.

où Λ_{mn} est un *opérateur* contenant les coordonnées et les moments conjugués des noyaux :

$$\Lambda_{mn} \stackrel{\text{déf}}{=} \int \prod_i \mathrm{d}q_i \sum_n \Phi_m^* \Big[\sum_k \frac{1}{M_k} (\vec{P}_k \Phi_n) \vec{P}_k + (T_\mathrm{N} \Phi_n) \Big] \ . \tag{28.26}$$

L'équation (28.25) peut s'écrire en notation de Dirac, en convenant que les *brackets* sous-entendent l'intégration sur tous les q_i et en considérant les coefficients $\chi_m(Q_j)$ comme des fonctions d'onde $\langle Q_j | \chi_m \rangle$:

$$[T_\mathrm{N}(Q_j) + E_m(Q_j)] \, |\chi_m\rangle \ = \ E \, |\chi_m\rangle - \sum_n \Big(\langle \Phi_m | T_\mathrm{N} | \Phi_n \rangle + \sum_k \langle \Phi_m | \frac{\vec{P}_k}{M_k} | \Phi_n \rangle \, . \, \vec{P}_k \Big) |\chi_n\rangle \ . \tag{28.27}$$

Les termes dans la sommation de droite n'ont pas d'interprétation simple, et ont pour origine la dépendance *paramétrique* des fonctions d'onde électroniques $\Phi_n(q_i \, ; \, Q_k)$ par rapport aux coordonnées nucléaires Q_k. Plus formellement, ceci s'écrit aussi :

$$\boxed{[T_\mathrm{N}(Q_j) + E_m(Q_j)] \, |\chi_m\rangle \ = \ E \, |\chi_m\rangle - \sum_n \Lambda_{mn} |\chi_n\rangle} \tag{28.28}$$

Clairement, Λ_{mn} constitue la matrice sur la base $\{|\Phi_n\rangle\}_n$ d'un certain opérateur représentant les couplages cinétiques entre les états électroniques obtenus dans la première étape de l'approximation de Born et Oppenheiner, et induisant *de facto* une dépendance entre les coefficients $\chi_m(Q_j)$ du développement (28.20).

Approximation adiabatique

Les équations (28.25) - (28.28) sont en principe exactes, dans l'hypothèse où les $\{\Phi_n\}$ forment une base complète. L'approximation dite adiabatique[7] consiste à laisser tomber purement et simplement tous les couplages Λ_{mn} et à ne retenir que :

$$[T_\mathrm{N} + E_m(Q_j)] \ \chi_m(Q_j) \ \simeq \ E \, \chi_m(Q_j) \ . \tag{28.29}$$

Alors, ou bien $\chi_m(Q_j) \equiv 0$, ou bien $\chi_m(Q_j)$ est l'une des fonctions propres de l'opérateur $T_\mathrm{N} + E_m(Q_j)$ figurant au premier membre de l'équation (28.29), le nombre quantique m étant fixé. Autrement dit, le développement (28.20) se réduit alors à un seul terme, celui associé au m choisi et la fonction d'onde totale est un simple produit :

$$\boxed{\Psi(q_i \, ; \, Q_j) \ \simeq \ \chi_m(Q_j) \, \Phi_m(q_i \, ; \, Q_j) \ \equiv \ \Psi_m^{\text{adiab}}(q_i \, ; \, Q_j)} \tag{28.30}$$

La justification détaillée de cette approximation n'est pas aisée. Elle signifie, physiquement, que le mouvement des électrons se fait à m donné, c'est-à-dire que le mouvement (lent, quasi-statique) des noyaux intervient de fait dans les états électroniques

[7]Tout comme à propos du théorème de Gell-Mann et Low [196] (voir chapitre 24, section 24.6), le qualificatif *adiabatique* évoque une variation lente de certains paramètres.

comme une simple variation *paramétrique* de chacun d'entre eux : on néglige de ce fait les transitions induites d'un état Φ_n à un autre état $\Phi_{n'}$; dit autrement, les électrons s'adaptent instantanément à la configuration lentement variable des noyaux, d'où le nom d'approximation adiabatique. D'un autre côté, la fonction nucléaire obéit à l'équation (28.29) où la dynamique des électrons n'apparaît plus explicitement (les coordonnées électroniques q_i n'y figurent pas, elles ont été moyennées sous la forme $\langle\Phi_m|H_0|\Phi_m\rangle$) : les noyaux voient en fait le mouvement *moyen* des électrons, en conformité avec l'idée que ceux-ci vont très vite.

La lenteur du mouvement nucléaire assure en général qu'une fréquence de vibration ω_{vibr} (et *a fortiori* une fréquence de rotation ω_{rot}) est toujours très petite devant ω_{el}, de sorte que les différents états électroniques, même munis de leur structure fine vibrationnelle et rotationnelle, restent le plus souvent bien séparés en énergie. *A contrario*, cette remarque permet de prévoir que l'approximation adiabatique est sûrement en défaut lorsque deux énergies E_n et E_m se rapprochent, voire en viennent à se croiser. Ces énergies étant paramétrées par les coordonnées nucléaires, ceci peut se produire quand le squelette moléculaire se déforme. Clairement, au voisinage d'un point de croisement, on ne peut faire l'économie d'un traitement plus raffiné du genre théorie de perturbation pour deux niveaux quasi-dégénérés, impliquant au moins un mélange des deux états Φ_n et Φ_m – à moins bien sûr que les éléments de matrice de Λ_{mn} soient nuls, pour des raisons de symétrie par exemple. Lorsque tel n'est pas le cas, on assiste comme toujours en pareil cas à un croisement de niveaux *évité* où la dégénérescence est levée[8]. Il en résulte des changements de propriétés physiques de la molécule (distorsion, modification des spectres, altération des mécanismes de réaction).

Hormis cette dernière situation, donnant lieu à des effets observables[9], la validation de l'approximation adiabatique est finalement fournie par l'explication remarquable des spectres moléculaires, tant dans le domaine UV, qu'IR ou micro-ondes.

28.1.3 Discussion

En résumé, dans l'approximation adiabatique, la fonction d'onde totale de la molécule est de la forme :

$$\Psi(q_i\,;\,Q_j) = \chi_m(Q_j)\,\Phi_m(q_i\,;\,Q_j)\ , \tag{28.31}$$

avec :

$$\boxed{H_0\Phi_m = E_m(Q_j)\Phi_m} \tag{28.32}$$

et :

$$\boxed{[T_{\text{N}} + E_m(Q_j)]\chi_m(Q_j) = E\chi_m(Q_j)} \tag{28.33}$$

La première équation, (28.32), est dite *équation électronique*, dont la fonction propre Φ_m est la fonction d'onde des électrons, les noyaux étant fixés ; E_m est la valeur propre

[8] Une situation analogue se présente pour un atome en champ magnétique, voir par exemple la situation étudiée dans le problème 23.4.11.

[9] Les plus connus portent le nom d'effet Renner, effet Jahn - Teller, effet Landau - Zener.

correspondante, qui dépend précisément des Q_j. La seconde équation, (28.33), est dite *équation nucléaire* ; la fonction propre χ_m décrit le mouvement des noyaux et satisfait une équation aux valeurs propres où la valeur propre électronique $E_m(Q_j)$ joue finalement le rôle d'une énergie potentielle *effective*, représentant les interactions entre noyaux *habillées* par le mouvement moyen des électrons[10]. L'énergie propre E figurant dans (28.33), en tant que valeur propre (scalaire), est l'énergie totale de la molécule, une fois prise la moyenne (au sens quantique) sur les variables électroniques. L'équation pour χ_m contient l'opérateur cinétique des noyaux, E représente donc bien l'énergie totale de la molécule (électrons et noyaux), dans l'approximation ainsi définie.

À propos du mouvement des électrons

Les états électroniques (Φ_m, E_m) dépendent des positions nucléaires, un fait qui va être illustré dans le cas d'une molécule diatomique AB. $\vec{R}$ désignant le vecteur $\overrightarrow{AB}$, ces états ne dépendent évidemment que de la distance $R = \mathrm{AB}$ (isotropie de l'espace). E_m est donc une fonction de R, qui a forcément une limite quand $R \to +\infty$ (sans que l'on sache précisément suivant quelle loi cette limite est atteinte) ; par ailleurs, elle tend vers l'infini quand $R \to 0$ (dans le cas contraire, la molécule serait instable aux courtes distances et tendrait à *collapser* sur elle-même), traduisant la résistance de chaque noyau (ou chaque cœur) à l'invasion par l'autre de son domaine personnel. Ceci étant acquis, deux cas sont essentiellement possibles :

- $E_m(R)$ présente un minimum pour une valeur R_0 (fig. 28.2, (a)), prouvant que la molécule est stable ; l'état Φ_m est alors dit *liant*. Si, en outre, E_m est la plus petite valeur propre électronique de (28.32), R_0 s'interprète comme la longueur de la liaison chimique AB. L'état ainsi décrit étant stable, il faut une énergie E_{diss} pour la dissocier en deux constituants séparés : c'est l'énergie de liaison. Si E_m n'est pas la plus petite énergie propre de (28.32), R_0 est la longueur de liaison (usuellement étirée) lorsque la molécule est dans un état excité stable vis-à-vis de la configuration nucléaire mais certes instable radiativement et/ou par conversion interne.
- E_m n'a pas de *minimum* (voir fig. 28.2, (b)) ; l'état Φ_m est alors dit *antiliant*. Ceci signifie que pour cet état électronique la molécule n'a pas d'état stable vis-à-vis de la configuration nucléaire et se dissocie spontanément. Ce type d'état est l'intermédiaire obligé de la photodissociation. Si E_m est la plus petite valeur propre de (28.32), l'absence de minimum signifie que cet édifice moléculaire est une espèce éphémère.

Ce sont les transitions entre états électroniques qui donnent les grands traits du spectre de la molécule dans l'optique ou l'UV (fig. 28.3 à gauche). Comme les mouvements nucléaires sont lents, la transition est quasi-verticale (les Q_j ne changent pas pendant

[10]On comprend que ce *rhabillage* est essentiel pour la stabilité de la molécule : des noyaux *nus* ne savent que se repousser violemment.

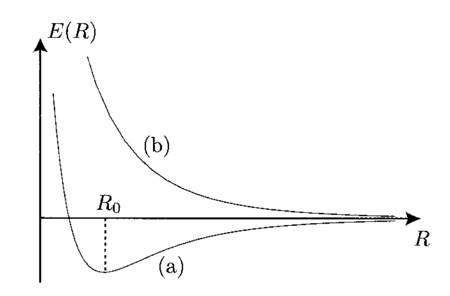


Figure 28.2: Variation de l'énergie électronique pour un état stable (a) ou instable (b) d'une molécle diatomique. Les valeurs des énergies en $R = +\infty$ ne sont pas forcément les mêmes.

l'absorption, flèche AB). Lorsque l'état excité a une durée de vie radiative assez longue, une relaxation essentiellement vibrationnelle peut alors se produire (passage de B à C). Pour les petites molécules, ceci suppose qu'elles ne sont pas pas isolées ; en phase vapeur, ce sont les collisions qui permettent cette relaxation – avec la complicité éventuelle des degrés de liberté de rotation, puisqu'il y a toujours un couplage entre vibration et rotation (quand la molécule vibre, les moments d'inertie changent). Après cette relaxation, la molécule à l'état (électronique) excité se retrouve dans la configuration d'équilibre caractérisée par Q_{j0}^* – où, généralement, elle a un peu *grossi*. L'émission radiative qui survient tôt ou tard[11] se fait à nouveau verticalement (flèche CD) et, après relaxation dans l'état fondamental (DA), la molécule retrouve la configuration Q_{j0} d'équilibre de cet état.

Il en résulte qu'en général[12] absorption et émission ne se produisent pas au même endroit, cette dernière donnant un pic décalé vers le *rouge* par rapport à l'absorption. En outre, les intensités sont en général assez différentes, pour la raison suivante : dans les éléments de matrice pilotant l'amplitude des transitions (donc l'intensité des raies), la partie électronique se factorise (la transition étant supposée dipolaire électrique) de la partie nucléaire ; dans cette dernière, on trouve le produit scalaire de deux oscillateurs différents, situés dans des niveaux de vibration différents, et appelé *facteur de Franck - Condon*. Le décalage entre absorption et émission (*Stokes shift*) est l'un des facteurs rendant difficile la *résonance optique* pour une molécule[13].

De la même façon, les états électroniques interviennent dans les processus de photodissociation : sous l'effet d'une irradiation optique ou UV, la molécule est portée de l'état fondamental (stable) à un état excité instable (fig. 28.3 à droite), qui se dissocie spontanément si la molécule n'est pas isolée.

[11] en l'absence d'autres processus de conversion interne.

[12] Bien sûr, on ne peut pas formellement exclure des coïncidences. Si la courbe inférieure est assez plate, et si les positions d'équilibre ne sont pas trop différentes, les deux transitions peuvent se produire à des énergies assez voisines.

[13] Pour une transition atomique, l'écart entre absorption et émission est essentiellement dû au recul du noyau, le plus souvent petit par rapport à la largeur naturelle – au contraire d'une transition nucléaire, voir Tome I p. 253 et problème 8.6.2.

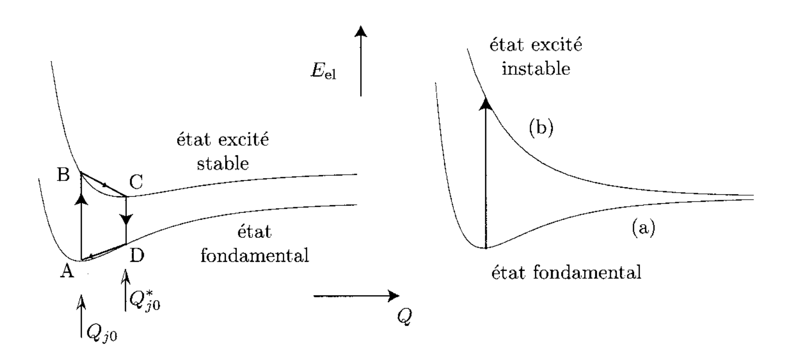


Figure 28.3: À gauche : Illustration schématique de l'absorption/émission d'une molécule ; l'absorption (AB) est suivie de l'émission (CD). Les petites flèches représentent la relaxation *via* les niveaux de rotation et vibration (non-représentés). À droite : photodissociation par excitation radiative d'un état stable (a) à un état instable (b).

À propos du mouvement des noyaux

En ce qui concerne les mouvements nucléaires, une molécule à N noyaux possède $3N$ degrés de liberté. 3 d'entre eux sont associés à la translation en bloc, 3 autres représentent la rotation. Ces mouvements réputés *solides* n'affectent pas le mouvement interne, de sorte que la valeur propre électronique E_m ne dépend en fait que de $3N-6$ coordonnées nucléaires[14]. Si R_j ($1 \leq j \leq 3N-6$ ou $3N-5$) désigne spécifiquement un degré de liberté de mouvement interne (associé à une certaine *déformation* de la molécule), on peut écrire, au voisinage d'un minimum absolu de E_m survenant pour les valeurs R_{j0} :

$$\Big[T_{\text{N}} + E_m(R_{j0}) + \frac{1}{2}\sum_{j,} k_{jj'}(R_j - R_{j0})\,(R_{j'} - R_{j'0}) + \ldots\Big]\,\chi_m(R_j) = E\chi_m(R_j) \ , \quad (28.34)$$

où les ... représentent des termes anharmoniques. En s'en tenant aux seuls termes écrits (approximation harmonique), on fait apparaître les petites vibrations des noyaux autour des positions d'équilibre. La forme quadratique se diagonalise en introduisant les bonnes combinaisons linéaires X_λ des coordonnées R_j, appelées modes normaux de vibration représentant des déformations *collectives* du squelette moléculaire. La détermination des vibrations normales est grandement facilitée par l'usage de la Théorie des groupes, plus particulièrement par la Théorie de la représentation linéaire des groupes : l'usage systématique de la symétrie permet, comme toujours, de prédiagonaliser la forme quadratique – voire de la diagonaliser complètement (voir l'exemple de la molécule plane X_3

[14]Pour une molécule linéaire, il n'y a en fait que $3N-5$ degrés de liberté internes. En effet, l'un des moments d'inertie est nul, et il ne reste que deux moments d'inertie donnant une énergie de rotation. L'un des degrés de liberté est non pertinent, c'est l'angle représentant la rotation autour de l'axe de la molécule. En pareil cas, il y a toujours 3 degrés de translation, mais seulement 2 de rotation.

Notons que ce décompte repose sur le présupposé que la molécule reste linéaire quoi qu'il arrive. Pour une molécule à plus de deux noyaux, ♡ – ◇ – ♠, ceci veut dire que toute déformation perpendiculaire à l'axe défini par la configuration d'équilibre, ♡ ╱ ◇ ╲ ♠, possède une constante de raideur très élevée.

étudié dans le problème 28.4.8). Par ailleurs, de nouvelles constantes de raideur K_λ apparaissent, qui s'expriment naturellement à l'aide des $k_{jj'}$.

Le terme T_{N} contient tous les moments conjugués (il y en a $3N$), 6 (ou 5) d'entre eux étant associés aux mouvements en bloc. Les 3 moments conjugués qui représentent la translation en bloc ne présentent pas d'intérêt ; les 3 (ou) 2 autres, associés à la rotation[15], contiennent les moments d'inertie, calculables *a priori* pour n'importe quelle configuration, pas seulement la configuration d'équilibre R_{j0}. Les $3N - 6(5)$ impliquant les modes normaux associés à la forme quadratique $\frac{1}{2}\sum_\lambda K_\lambda X_\lambda^2$ (avec K_λ strictement positif) donnent $3N - 6(5)$ oscillateurs harmoniques indépendants.

La variation des moments d'inertie avec l'état de vibration des liaisons induit un couplage entre la vibration et la rotation. Cet effet est souvent assez petit, et peut être omis dans un premier temps ; lorsqu'il est mis de côté, la rotation se découple de la vibration, les moments d'inertie étant alors naturellement définis relativement à la configuration d'équilibre. Dans ces conditions, l'énergie totale de la molécule[16] se décompose en une *somme* de trois termes, la simple somme résultant du découplage des degrés de liberté :

$$\boxed{E = E_m(R_{j0}) + E_{\mathrm{vibr}} + E_{\mathrm{rot}}} \tag{28.35}$$

$E_m(R_{j0})$ est la valeur de l'énergie électronique calculée dans la configuration d'équilibre ; E_{vibr} est l'énergie de l'ensemble des oscillateurs au-dessus du minimum d'énergie électronique ; enfin, E_{rot} est l'énergie de rotation, calculée avec les noyaux occupant leurs positions d'équilibre et donc *ipso facto* découplée de la vibration. Les variations de ces trois termes se produisent sur des échelles d'énergie très différentes, respectivement UV ou optique, IR et micro-ondes. La discussion détaillée des deux types de mouvement et leurs conséquences observables sont l'objet de la section 28.3.

28.2 Structure électronique des molécules. Nature physique de la liaison chimique

Dans la section 28.1, on a vu comment il est possible de séparer les différents degrés de liberté d'un édifice complexe comme une molécule. Tout naturellement, l'étude des propriétés moléculaires commence par l'analyse des états électroniques, qui sont les plus énergétiques et qui, pour une grande part, déterminent les propriétés chimiques. Sur un plan conceptuel, c'est bien la Mécanique quantique qui fournit une description rationnelle de l'origine de la liaison chimique et, par là, permet de comprendre la stabilité des molécules. Il s'agit ici d'exposer les idées principales – en prenant pour exemples les édifices moléculaires les plus simples – sans rentrer dans une zoologie quasi-inépuisable. Les plus simples, et aussi les plus exemplaires : le traitement élémentaire de la molécule de benzène (sous-section 28.2.4) n'est qu'une préfiguration de l'approximation des liaisons fortes (chapitre 30, sous-section 30.4.2) ; quant au polyacétylène (sous-section 28.2.5), il

[15] décrite à l'aide des angles d'Euler (voir [15], p. 447)
[16] dans le repère de son centre de masse.

anticipe sur l'instabilité de Peierls dont il sera question dans le chapitre 30. En un sens, et tout naturellement, la molécule est à mi-chemin entre l'atome et le solide, et c'est à propos de physique moléculaire que l'on retrouve une sorte de synthèse des notions physiques qui s'imposent dans un cas et dans l'autre.

28.2.1 L'ion moléculaire H_2^+

La molécule la plus simple – au sens d'assemblage de plusieurs noyaux – est formée de deux protons avec juste ce qu'il faut d'électrons (en fait un seul) pour fabriquer un ciment susceptible de s'opposer à la répulsion électrostatique des deux protons nus. *A priori*, rien ne permet d'affirmer qu'un tel système existe à l'état stable : l'électron unique est tiraillé entre les deux protons, ce qui introduit une interaction effective *attractive* entre ces derniers. La question de savoir si cette attraction est suffisante pour lutter contre la répulsion coulombienne entre protons ne peut être tranchée que par une étude *quantitative*, donnée ci-dessous dans le cadre de l'approximation adiabatique ; ici, la seule coordonnée nucléaire pertinente est la distance R entre les protons.

En tout état de cause, on voit de suite que c'est effectivement le *partage* d'un électron entre les protons qui est susceptible de produire une liaison stable, prototype de la liaison dite covalente. On rencontre ici un exemple du schéma universel représentant toute interaction physique : deux particules en interaction sont deux particules échangeant entre elles d'autres particules (ce sont ces dernières qui véhiculent l'échange d'"information" se traduisant par une interaction). Ici, les allers-retours de l'électron d'un proton à l'autre induisent une interaction *effective* attractive entre ceux-ci, quelle que soit leur distance R ; complétée par la répulsion $\frac{e'^2}{R}$, elle constitue l'énergie[17] $E(R)$ figurant dans l'équation nucléaire au titre d'une énergie potentielle. C'est d'un bilan quantitatif entre ces deux tendances antagonistes que peut, éventuellement, résulter un édifice stable. On note que si un tel état existe, l'interaction globale ne ressemble pas du tout à l'interaction nue des deux protons – qui se repoussent suivant la loi de Coulomb (*renormalisation* des interactions) : c'est bien l'échange (c'est-à-dire le partage) de l'électron qui peut finalement créer une paire liée proton - proton[18]. L'électron voyageur constitue bel et bien le *ciment* moléculaire.

[17]Notée $E(Q_j)$ dans la sous-section précédente, l'énergie électronique ne dépend ici que d'une seule variable, R.

[18]*Mutatis mutandis*, le même scénario peut se produire dans un métal à basse température pour donner lieu à la supraconductivité : alors, ce sont les phonons du réseau (vibrations quantifiées) qui servent d'intermédiaires et permettent à deux électrons de former une paire liée électron - électron (voir chapitre 32, sous-section 32.5.2).

Autre exemple de mise en évidence explicite de l'*échange* des particules qui sont les *vecteurs* d'une interaction : les photons, responsables des forces de van der Waals entre deux objets électriquement neutres (voir chapitre 25, section 25.4).

Solution exacte

S'agissant d'obtenir les états électroniques tels que les définit l'approximation de Born - Oppenheimer, il faut résoudre l'équation aux valeurs propres où les noyaux sont fixes, séparés par la distance R, soit trouver les solutions de :

$$H_0\Phi = E\,\Phi \ , \tag{28.36}$$

H_0 étant le Hamiltonien électrostatique à noyaux fixes :

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r_1} - \frac{e'^2}{r_2} + \frac{e'^2}{R} \ . \tag{28.37}$$

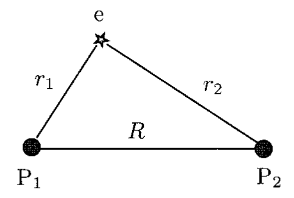


Figure 28.4: L'ion moléculaire H_2^+.

Φ est une fonction de la position $\vec{r}$ de l'électron (exprimée d'une façon ou d'une autre), paramétrée par la distance[19] internucléaire R, soit $\Phi(\vec{r};\, R)$. r_1 et r_2 désignent les distances de l'unique électron aux deux protons (voir fig. 28.4) ; noter que les r_i contiennent la distance R. Dans l'expression (28.37), $\frac{e'^2}{R}$ est une simple constante additive, qu'il ne faut cependant pas omettre : elle est une contribution importante à l'énergie totale qui joue un rôle décisif pour l'apparition d'un minimum d'énergie quand la distance internucléaire varie. L'étape importante est donc de fait la résolution de[20] :

$$\Big(\frac{\vec{p}^2}{2m} - \frac{e'^2}{r_1} - \frac{e'^2}{r_2}\Big)\Phi(\vec{r};\, R) = E_{\mathrm{el}}(R)\Phi(\vec{r};\, R) \ , \tag{28.38}$$

étant entendu que l'énergie E définie en (28.36) est :

$$\boxed{E(R) = E_{\mathrm{el}}(R) + \frac{e'^2}{R}} \tag{28.39}$$

Il se trouve que pour ce système très simple, il est possible de résoudre exactement l'équation aux valeurs propres, comme l'a montré[21] Burrau [219]. Le calcul complet étant

[19]Évidemment, les états propres de (28.36) ne dépendent pas de l'orientation de l'axe portant P_1 et P_2.

[20]Ici, il n'y a qu'un électron : la fonction Φ est donc *de facto* une orbitale ; comme elle est décrit un électron dans un assemblage de plusieurs noyaux, on l'appelle *orbitale moléculaire*. Dans (28.38), $\vec{r}$ est (par exemple) le rayon-vecteur de l'électron en prenant l'origine au milieu de P_1P_2.

[21]Burrau était un spécialiste du problème des trois corps en Mécanique (voir note 18, p. 1051).

très technique, on s'en tiendra à quelques éléments purement descriptifs, donnés à titre d'esquisse de celui-ci.

Le problème a clairement une symétrie cylindrique ; l'angle ϕ entre le plan contenant l'électron et les noyaux et un plan de référence arbitraire est une variable cyclique (elle est absente de H_0). De fait, si $\vec{l}$ désigne le moment cinétique de l'électron par rapport à un point quelconque de l'axe internucléaire (choisi comme axe Oz), alors l_z est une constante du mouvement :

$$[H_0, l_z] = 0 \ , \tag{28.40}$$

relation traduisant la symétrie cylindrique du système, et la variable ϕ se sépare. Les fonctions propres de $l_z = -\mathrm{i}\hbar\frac{\partial}{\partial\phi}$ sont de la forme $\mathrm{e}^{\mathrm{i}m\phi}$, avec m entier comme l'exige la théorie générale du moment cinétique orbital dès que les fonctions propres doivent être monovaluées. L'usage a consacré la notation σ, π, δ,..., suivant que $|m| = 0, 1, 2, \ldots$, en réminiscence de la notation s, p, d, ... pour les états atomiques. Par ailleurs, H_0 étant invariant par renversement du temps[22] – opération qui inverse le moment cinétique – on peut affirmer d'emblée que l'énergie propre associée à Φ, qui est évidemment fonction de m puisqu'il n'y a pas de symétrie sphérique comme dans un atome, ne dépend en fait que de $|m|$, pas du signe de m.

Compte tenu de la géométrie du système, les coordonnées elliptiques s'imposent :

$$\xi \stackrel{\text{déf}}{=} \frac{r_1 + r_2}{R} \ , \quad 1 \leq \xi < +\infty \ , \qquad \mu \stackrel{\text{déf}}{=} \frac{r_1 - r_2}{R} \ , \quad -1 \leq \mu \leq +1 \ ; \tag{28.41}$$

ce choix étant fait, tous les états propres Φ peuvent être cherchés sous la forme[23] :

$$\Phi(\xi, \mu, \phi) = F(\xi, \mu)\,\mathrm{e}^{\mathrm{i}m\phi} \qquad (m \in \mathbb{Z}) \ , \tag{28.42}$$

où F est une fonction à déterminer. Le Laplacien en coordonnées elliptiques est :

$$\Delta = \frac{4}{R^2(\xi^2 - \mu^2)}\Big[(\xi^2 - 1)\frac{\partial^2}{\partial\xi^2} + 2\xi\frac{\partial}{\partial\xi} - (\mu^2 - 1)\frac{\partial^2}{\partial\mu^2} - 2\mu\frac{\partial}{\partial\mu} + \left(\frac{1}{\xi^2 - 1} - \frac{1}{\mu^2 - 1}\right)\frac{\partial^2}{\partial\phi^2}\Big] \ ; \tag{28.43}$$

la séparation (28.42) remplace simplement $\frac{\partial^2}{\partial\phi^2}$ par $-m^2$. L'équation aux fonctions propres pour $F(\xi, \mu)$ s'écrit ainsi :

$$-\frac{\hbar^2}{mR^2}\Big[(\xi^2 - 1)\frac{\partial^2}{\partial\xi^2} + 2\xi\frac{\partial}{\partial\xi} - (\mu^2 - 1)\frac{\partial^2}{\partial\mu^2} - 2\mu\frac{\partial}{\partial\mu} - m^2\Big(\frac{1}{\xi^2 - 1} - \frac{1}{\mu^2 - 1}\Big)\Big] F(\xi, \mu) - \frac{e'^2}{R}\Big[2\xi - \frac{1}{2}(\xi^2 - \mu^2)\Big] F(\xi, \mu) = \frac{1}{2}E\,(\xi^2 - \mu^2)\,F(\xi, \mu) \ ; \tag{28.44}$$

---

[22] H_0 ne contient que le *carré* de la vitesse.

[23] La fonction F est paramétrée par m, nombre qui apparaît explicitement dans les équations pour les fonctions f et g introduites ci-dessous ; cette dépendance n'est toutefois pas explicitée pour ne pas alourdir les notations.

visiblement, les deux variables ξ et μ se séparent aussi dans (28.44) ; sans perte de généralité, on peut donc poser :

$$F(\xi,\,\mu) = f(\xi)\,g(\mu) \ ; \tag{28.45}$$

le report dans (28.44) montre que chacune des fonctions f et g satisfait une équation différentielle associée de Legendre (voir éq. (18.95)) [220], paramétrée par m^2 (et R), dont les solutions sont bien connues et répertoriées. Comme toujours en Mécanique quantique, il convient de faire le tri parmi toutes les solutions possibles, pour ne garder que celles qui sont physiquement acceptables. S'agissant d'obtenir les états liés, donc normalisables, cette condition agit comme un crible et ne retient qu'un nombre infini dénombrable de solutions associés à des énergies discrètes (quantifiées) $\{E_{n,\,\text{el}}(R)\}_{n\in\mathbb{N}^*}$, toujours dépendantes de R, qui est contenu dans la définition (28.41) des coordonnées elliptiques.

Cette solution exacte du problème ne donne pas d'expressions simples, ni pour les fonctions f et g, qui sont des séries de polynômes de Legendre[24] – l'existence d'une série en l traduisant le fait que le carré du moment cinétique orbital n'est pas une constante du mouvement (brisure de la symétrie sphérique) –, ni pour les énergies $E_{n,\,\text{el}}(R)$. On peut néanmoins tracer numériquement la variation de ces dernières en fonction de la distance internucléaire R. Les énergies $E_n(R)$ définies par (28.39) pour les premiers états sont données dans l'ouvrage de Slater [215] et reportées sur la figure 28.5. Il apparaît deux types d'états, stables et instables (ce sont les états *liants* et *antiliants* déjà introduits), pouvant d'ailleurs se croiser[25]. Les deux états de plus basse énergie forment une paire liante/antiliante, ce que l'on peut comprendre par un argument peu rigoureux de théorie de perturbation : si $R = +\infty$, chaque niveau est dégénéré deux fois, l'électron pouvant se trouver soit sur un proton, soit sur l'autre. Quel que soit R fini[26], les deux éléments diagonaux de H sont encore égaux par symétrie, mais l'élément non-diagonal n'est plus nul : c'est lui qui va lever la dégénérescence, à peu près symétriquement tant que R n'est pas trop petit ; on obtient deux niveaux distincts de nature opposée, formés avec les deux combinaisons linéaires $\pm$ des deux seuls états atomiques introduits.

La divergence de $E_n(R)$ à l'origine provient exclusivement de $\frac{e'^2}{R}$. En ce qui concerne les $E_{n,\,\text{el}}(R)$, leurs limites sont faciles à trouver :

- $R \to 0$: on a un ion He^+ ($Z = 2$), dont les énergies sont $-\frac{2^2 me'^4}{2n^2\hbar^2} = 4E_{n,\,\text{hydrogène}}$ soit $-4\frac{E_\text{I}}{n^2}$ où :

$$E_\text{I} = -\frac{\hbar^2}{2ma_0^2} = -\frac{e'^2}{2a_0} \simeq -13,61 \text{ eV} \ , \tag{28.46}$$

[24] Par exemple, $g(\mu) = \sum_{l\geq m} \alpha_l P_l^{l-m}(\mu)$, où α_l est un certain coefficient.

[25] au voisinage des points de croisement, l'approximation adiabatique doit être reconsidérée – et peut se révéler d'ailleurs encore valide si, pour des raisons de symétrie par exemple, les deux états restent découplés.

[26] mais pas trop petit afin que l'on puisse s'en tenir aux deux états 1s des atomes séparés. Aux petites distances, c'est l'ensemble des états atomiques qu'il faut considérer, *a priori*.

d'où :

$$\lim_{R \to 0} E_{n,\,\mathrm{el}}(R) = -4\frac{E_\mathrm{I}}{n^2} \ . \tag{28.47}$$

- $R \to \infty$: on se retrouve avec un proton nu et un atome d'hydrogène, système d'énergie $-\frac{E_\mathrm{I}}{n^2}$.

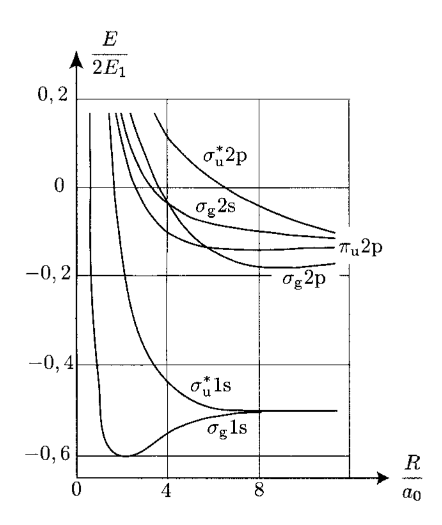


Figure 28.5: Variation de l'énergie totale $E_n(R)$ pour les premiers états de l'ion H_2^+. $\frac{R}{a_0}$ est porté en abscisse, $\frac{E}{2E_\mathrm{I}}$ est en ordonnée ; $E_\mathrm{I} = \frac{1}{2}\alpha^2 mc^2 \simeq 13,61$ eV est l'unité atomique d'énergie, aussi appelée Rydberg (on définit aussi le *Hartree*, égal à 2 Rydbergs). Noter l'existence de croisements, pour lesquels il convient de réexaminer au cas par cas l'approximation adiabatique.

Pour l'énergie totale (28.39) (les noyaux étant toujours fixes), on a donc :

$$R \to 0 \ : \ E_n(R) \simeq -4\frac{E_\mathrm{I}}{n^2} + \frac{e'^2}{R} \simeq \frac{e'^2}{R} \ , \tag{28.48}$$

$$R \to +\infty \ : \ E_n(R) \simeq -\frac{E_\mathrm{I}}{n^2} + \frac{e'^2}{R} \simeq -\frac{E_\mathrm{I}}{n^2} \ . \tag{28.49}$$

Pour l'état fondamental de H_2^+, les différentes contributions à l'énergie se combinent pour donner effectivement un état stable, et comme $\frac{e'^2}{R}$ diverge en $R = 0$, l'existence d'un tel état liant est finalement conditionné par la décroissance assez rapide de $E_{n,\,\mathrm{el}}(R)$ quand R décroît (voir plus loin la discussion physique). Quantitativement, le minimum est ici assez plat et l'énergie de liaison E_diss est corrélativement faible ; les valeurs numériques sont :

$$\boxed{R_0 = 2a_0 \simeq 1,06 \text{ Å} \qquad E_\mathrm{diss} \simeq 2,79 \text{ eV} \simeq 64\,\mathrm{kcal/mole} \simeq 0,17\,E_\mathrm{I}} \tag{28.50}$$

L'énergie de dissociation ne représente donc que 17% de l'énergie totale au minimum. Au total, la molécule est assez "molle" (fragile), ce qui se comprend bien puisqu'il n'y a qu'un seul électron pour assurer la liaison.

La notation traditionnelle des différents états fait référence à leurs propriétés de symétrie. En particulier, H_0 est invariant dans l'inversion d'espace Π (parité) ; cette opération change ϕ en $\phi + \pi$ sans modifier ni r_1, ni r_2 et n'affecte donc visiblement pas H_0 ($[\Pi,\, H_0] = 0$). En l'absence de dégénérescence, les états propres sont donc soit pairs (notés g, pour *gerade*), soit impairs (notés u, pour *ungerade*) dans l'inversion :

$$\Pi\Phi_{\rm g} = \Phi_{\rm g} \ , \qquad \Pi\Phi_{\rm u} = -\Phi_{\rm u} \ . \tag{28.51}$$

Un autre élément de symétrie physiquement utile à préciser est le comportement des fonctions propres dans une réflexion par rapport au plan médiateur des deux protons P_1 et P_2, puisqu'une fonction impaire dans cette opération est nulle partout dans ce plan (on s'attend donc à ce qu'elle soit antiliante). Cette propriété de symétrie n'est pas indépendante des précédentes puisqu'elle est le produit d'une rotation d'un angle π autour de l'axe internucléaire (qui donne $(-1)^m$ par le facteur $e^{im\phi}$) et d'une inversion d'espace. Le comportement des fonctions propres sous l'effet de la symétrie-miroir est donc :

- pas de changement de signe pour les états g ayant m pair et les états u avec m impair,
- changement de signe pour les états g ayant m impair et les états u avec m pair.

Malgré la redondance, l'usage est d'ajouter un astérisque[27] pour les états impairs dans la symétrie-miroir ($1\sigma_{\rm u}^*$, $2\pi_{\rm g}^*$, ...) : l'astérisque permet ainsi d'identifier à l'œil les orbitales antiliantes qui, par symétrie, s'annulent dans le plan médiateur ; pour une orbitale antiliante, la densité électronique est *nulle* au milieu de la liaison.

28.2.2 Méthodologie pour une description approchée

La méthode exposée ci-dessus est d'une part déjà complexe – alors qu'il s'agit de la molécule la plus simple –, d'autre part n'est pas généralisable au cas d'édifices plus complexes[28]. Il faut donc trouver d'autres méthodes de description (approchée, cette fois), qui constituent un véritable arsenal, quelque peu technologique, mais dont les idées fondamentales peuvent déjà s'énoncer à propos de H_2^+.

L'une de ces méthodes (dite LCAO, pour *Linear Combination of Atomic Orbitals*) consiste à définir une base de fonctions d'origine atomique pour exprimer, sous la forme

[27] Il n'existe pas d'états du type $\sigma_{\rm g}^*$, par exemple.

[28] On ne voit pas bien ce que viendraient faire les coordonnées elliptiques dans une molécule à $M > 2$ noyaux...

d'un développement linéaire, les orbitales moléculaires d'un électron[29]. L'idée physique sous-jacente est élémentaire. Chaque proton (ou noyau) offre à l'électron une infinité d'états liés et non-liés (états de l'atome d'hydrogène), mais il n'y a pas lieu de privilégier l'un de ces ensembles d'états – dont la réunion forme d'ailleurs un système surcomplet si l'on incorpore les états de diffusion. S'agissant de décrire le fondamental de l'ion et/ou ses premiers états excités, on choisit naturellement d'introduire les états atomiques de plus basse énergie. N'en retenant que deux – pour l'exemple, et pouvoir faire le calcul à la main –, on considère ainsi les deux états 1s normalisés, centrés chacun sur un proton, soit la fonction $\psi_{nlm} = \psi_{100} \equiv \phi_1$ centrée sur P_1, et sa jumelle centrée sur P_2 :

$$\phi_2(\vec{r}) = \psi_{100}(\vec{r} - \vec{R}) \ , \tag{28.52}$$

image de ϕ_1 par la translation $\mathcal{T}(\vec{R})$ (on prend l'origine du repère sur P_1). Ces deux fonctions (réelles) ne sont pas orthogonales ; soit S leur produit scalaire :

$$\int_{\mathbb{R}^3} \phi_1(\vec{r})\phi_2(\vec{r})\,\mathrm{d}^3r = S \ , \tag{28.53}$$

qui est bien sûr une fonction décroissante de R, en gros de façon exponentielle. D'ailleurs, en écrivant la positivité du produit scalaire $\langle\phi_1 + \lambda\phi_2|\phi_1 + \lambda\phi_2\rangle \ \forall\,\lambda$, on voit facilement que $S^2 \leq 1$; ici, S est visiblement positif, de sorte que $0 \leq S \leq 1$.

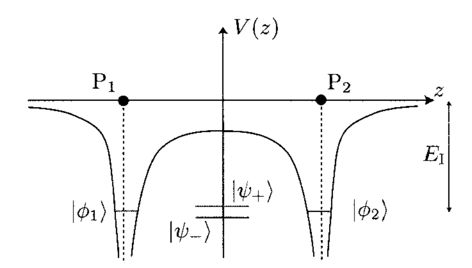


Figure 28.6: Double puits de potentiel pour l'électron dans H_2^+. Le passage par effet tunnel d'un proton à l'autre se fait à une fréquence donnée par la différence d'énergie entre les deux états $|\psi_\pm\rangle$ qui sont, par nature, *délocalisés* (étendus) au contraire des deux états atomiques $|\phi_i\rangle$

Le schéma le plus simple consiste donc à s'en tenir à ces deux fonctions, arguant du fait que les autres états sont nettement plus hauts en énergie ; la solution ainsi obtenue sera évidemment une solution approchée. Avant de commencer ce calcul, il est utile de donner une image physique du choix qui est fait ici. En gros, quand la distance R est très grande, l'unique électron est soit sur un proton, soit sur l'autre, et, de fait, les deux protons fabriquent un double puits de potentiel $V(r)$ qui est la simple somme des puits de chacun d'entre eux (voir fig. 28.6). En fait, en pareille situation, l'électron peut passer par effet tunnel d'un côté à l'autre du puits ; tant que la distance est très grande, la

[29]Pour une molécule à plusieurs électrons, on utilisera les orbitales moléculaires pour contruire un (des) déterminant(s) de Slater, tout comme on utilise des orbitales atomiques pour contruire les états antisymétriques d'un atome polyélectronique.

probabilité de saut est exponentiellement petite. Elle croît très vite quand la distance R diminue. La description minimale est bien de prendre une fonction localisée sur chaque proton : si elle peut sembler quelque peu simpliste, elle n'en retient pas moins l'essentiel du phénomène physique sous-jacent. On écrit ainsi, pour la fonction électronique (mais ici, c'est une simple orbitale moléculaire puisqu'il n'y a qu'un électron) :

$$\Phi = c_1\phi_1 + c_2\phi_2 \ , \tag{28.54}$$

et la question est maintenant de trouver les deux coefficients c_1 et c_2. En reportant dans l'équation aux valeurs propres (28.36), en multipliant à gauche par $\langle\phi_1|$ et par $\langle\phi_2|$, on obtient le système[30] :

$$\left\{ \begin{array}{l} \langle\phi_1|H_0|\phi_1\rangle\, c_1 + \langle\phi_1|H_0|\phi_2\rangle\, c_2 = E(c_1 + Sc_2) \\ \langle\phi_2|H_0|\phi_1\rangle\, c_1 + \langle\phi_2|H_0|\phi_2\rangle\, c_2 = E(Sc_1 + c_2) \end{array} \right. . \tag{28.55}$$

En posant :

$$\langle\phi_1|H_0|\phi_1\rangle = \langle\phi_2|H_0|\phi_2\rangle \stackrel{\text{déf}}{=} \varepsilon \ , \qquad \langle\phi_1|H_0|\phi_2\rangle = \langle\phi_2|H_0|\phi_1\rangle \stackrel{\text{déf}}{=} v \ , \tag{28.56}$$

on obtient :

$$\left\{ \begin{array}{l} (\varepsilon - E)\, c_1 + (v - E\, S)\, c_2 = 0 \\ (v - E\, S)\, c_1 + (\varepsilon - E)\, c_2 = 0 \end{array} \right. . \tag{28.57}$$

Ce système n'a de solution non-nulle que si son déterminant est nul :

$$\left| \begin{array}{cc} \varepsilon - E & v - E\, S \\ v - E\, S & \varepsilon - E \end{array} \right| = 0 \iff \varepsilon - E = \pm(v - ES) \iff E = \frac{\varepsilon \pm v}{1 \pm S} \stackrel{\text{déf}}{=} E_\mp \ . \tag{28.58}$$

Pour E_+, on a $\varepsilon - E_+ = (v - E_+S)$, donc $c_1 + c_2 = 0$, soit $c_1 = -c_2$, l'état propre est de la forme $\phi_1 - \phi_2$ (état antiliant, densité nulle dans le plan médiateur) ; au contraire, pour E_-, on a $\varepsilon - E_- = -(v - E_-S)$, donc $c_1 - c_2 = 0$, soit $c_1 = c_2$: l'état propre est de la forme $\phi_1 + \phi_2$ (état liant, accumulation de densité entre les protons). Les deux fonctions propres (approchées) ainsi obtenues, une fois normalisées, sont donc :

$$\boxed{\Phi_{E_\pm} = \frac{1}{\sqrt{2(1 \mp S)}}\,(\phi_1 \mp \phi_2) \ , \qquad E_-(R) < E_+(R)} \tag{28.59}$$

Les deux énergies $E_\pm(R)$ varient comme indiqué sur la figure 28.7. Ce schéma très simple rend donc compte qualitativement de l'existence d'un fondamental stable. Bien sûr, on a l'inégalité :

$$E_-(R) > E_{n=1}(R) \ , \tag{28.60}$$

[30] Les deux orbitales atomiques ϕ_1 et ϕ_2 n'étant pas orthogonales, une quantité du type $\langle\phi_j|H_0|\phi_{j'}\rangle$ *n'est pas* l'élément de matrice $(H_0)_{jj'}$ de H_0 entre les deux états $|\phi_j\rangle$ et $|\phi_{j'}\rangle$ (voir Tome I, p. 446, le commentaire suivant l'équation (13.64)). En termes de ces quantités, l'équation donnant les valeurs propres E est $\text{Det}\big(\langle\phi_j|H_0|\phi_{j'}\rangle - E\langle\phi_j|\phi_{j'}\rangle\big) = 0$, équivalente à $\text{Det}\big((H_0)_{jj'} - E\delta_{jj'}\big) = 0$.

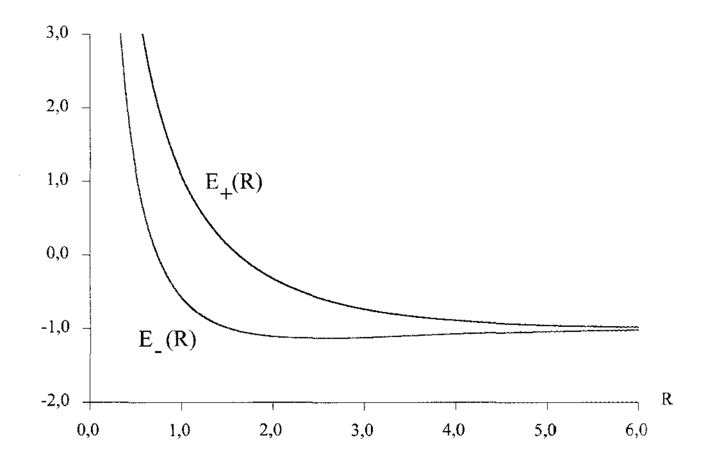


Figure 28.7: Variation de l'énergie totale pour les deux orbitales moléculaires liante et antiliante $E_{\mp}(R)$ (voir (28.58)). R est compté en unités a_0 et les énergies en unités $E_{\rm I}$; au minimum de $E_{-}(R)$, la différence d'énergie est voisine de 10 eV.

où $E_{n=1}(R)$ est l'énergie fondamentale exacte obtenue dans la section 28.2.1, conformément au Principe variationnel.

La figure 28.8 donne par comparaison les deux énergies électroniques correspondantes $E_{\rm el}(R)$, dont les limites en $R = 0$ ne sont pas (et n'ont aucune raison d'être) les mêmes que pour la fonction exacte (comparer avec (28.47) – avec quel n d'ailleurs ?). Pour l'état liant, le minimum est en $R = 0$, pour l'état antiliant, le minimum (très plat !) est pour un R'_0 fini. Dans les deux cas, l'interaction entre les deux protons médiée par l'électron est bien de fait attractive ; le terme répulsif nu $\frac{e'^2}{R}$ est assez fort pour faire disparaître le minimum très plat de la plus grande énergie, mais pas assez pour supprimer celui de l'autre.

Cela vaut la peine de dire un mot sur les quantités S, ε et v (pour plus de détails, voir [25], complément $G_{\rm XI}$, et problème 28.4.2 ci-après). Tout d'abord, on trouve sans difficulté (on passe en coordonnées elliptiques) :

$$S = \mathrm{e}^{-\rho}\,(1 + \rho + \frac{\rho^2}{3}) \qquad (\rho = \frac{R}{a_0}) \ ; \tag{28.61}$$

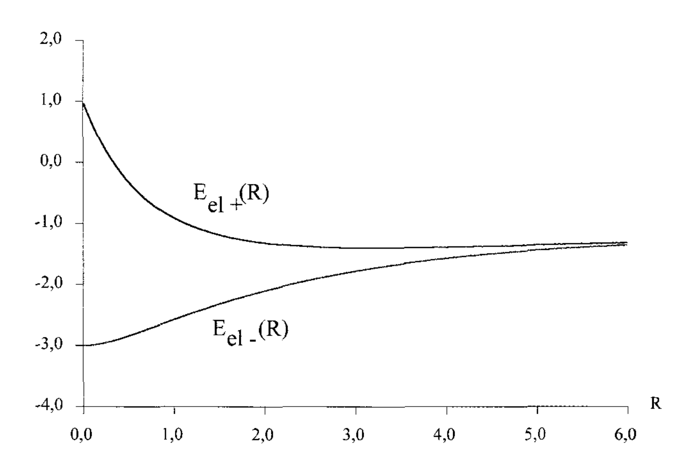


Figure 28.8: Variation de l'énergie électronique $E_{\text{el}\pm} = E_\pm - \frac{e'^2}{R}$ pour les deux orbitales moléculaires liante et antiliante $E_\mp(R)$ (voir (28.58)). R est compté en unités a_0 et les énergies en unités E_I.

par ailleurs, en raison de la définition[31] (28.56) :

$$\varepsilon = \langle\phi_1|\frac{\vec{p}^2}{2m} - \frac{e'^2}{r_1} - \frac{e'^2}{r_2}|\phi_1\rangle + \frac{e'^2}{R} = -E_\text{I} - \langle\phi_1|\frac{e'^2}{r_2}|\phi_1\rangle + \frac{e'^2}{R} \ ; \tag{28.62}$$

on trouve :

$$\langle\phi_1|\frac{e'^2}{r_2}|\phi_1\rangle = \frac{2}{\rho}\left[1 - \mathrm{e}^{-2\rho}(1+\rho)\right] E_\text{I} \ , \tag{28.63}$$

d'où :

$$\varepsilon = -E_\text{I} - \frac{2}{\rho}\left[1 - \mathrm{e}^{-2\rho}(1+\rho)\right] E_\text{I} + \frac{e'^2}{R} = -E_\text{I} + \frac{2}{\rho}\mathrm{e}^{-2\rho}(1+\rho)\, E_\text{I} \ . \tag{28.64}$$

Enfin, l'intégrale v est :

$$v = \langle\phi_1|H_0|\phi_2\rangle = -S\,E_\text{I} - \langle\phi_1|\frac{e'^2}{r_2}|\phi_2\rangle + S\,\frac{e'^2}{R} \ ; \tag{28.65}$$

un calcul simple donne l'expression finale de v :

$$v = -S\,E_\text{I} - 2\,\mathrm{e}^{-\rho}(1+\rho)\,E_\text{I} + S\,\frac{e'^2}{R} \ . \tag{28.66}$$

En définitive, c'est essentiellement[32] l'intégrale v (parfois appelée *intégrale de résonance*) qui offre la possibilité à l'électron de sauter par effet-tunnel d'un proton

[31] Se souvenir que $\phi_1 \equiv \psi_{100}$ est propre de la somme des deux premiers termes apparaissant dans (28.37).

[32] Si v était mis à zéro, on aurait encore $E_- < E_+$ à cause du dénominateur $1 \pm S$, mais la stabilisation serait bien moindre, et ne résulterait que du recouvrement des orbitales atomiques.

à l'autre, à une fréquence[33] $\nu = h^{-1}(E_+ - E_-)$: si on imagine l'électron "préparé" dans l'état purement atomique $\Psi(t = 0) = \phi_1$ centré sur P_1, l'état à l'instant t est $\Psi(t) = \cos 2\pi\nu t\, \phi_1 + \mathrm{i} \sin 2\pi\nu t\, \phi_2$. Les états ϕ_i sont localisés autour de chaque proton ; au contraire, les deux orbitales moléculaires sont *délocalisées* au sens où elles donnent une densité notable sur tout l'espace moléculaire. Cette délocalisation induite par la proximité de ces derniers produit un état d'énergie plus basse que $-E_\mathrm{I}$; on parle alors de stabilisation par résonance[34]. On retrouvera ci-dessous le fait que c'est bien le *mouvement* de l'électron sur tout l'espace offert qui est à l'origine d'une liaison stable (l'énergie cinétique de celui-ci *augmente* quand la liaison se forme, voir section 28.2.6).

Quantitativement, les résultats ne sont pas très fameux, mais la méthode peut, cette fois, être améliorée de façon évidente :

- en incluant dans les orbitales atomiques des paramètres variationnels (par exemple, une charge nucléaire ajustable) et en minimisant l'énergie liante, conformément au Principe variationnel. La charge effective obtenue n'est pas interprétable, au contraire d'un système à plusieurs électrons, par un effet d'écran ; c'est en fait un simple facteur d'échelle, comme on l'a vu à propos de l'atome d'hélium.
- en élargissant la base par l'introduction d'autres orbitales atomiques, ce qui revient à travailler dans un espace vectoriel plus vaste. D'une façon générale, si on introduit N orbitales atomiques, on obtiendra autant d'orbitales moléculaires qui seront des approximations pour les fonctions d'onde des N premiers états (y compris le fondamental) électroniques Φ_n. Les énergies correspondantes $E_n^{(N)}(R)$ $(1 \leq n \leq N)$ sont les énergies approchées, dont on espère qu'elles convergent vite vers leurs limites exactes quand N devient assez grand. Elles obéissent à certaines relations d'ordre quand on compare les résultats d'un premier calcul avec N orbitales atomiques, et ceux d'un second effectué en ajoutant deux orbitales atomiques à la base introduite dans le premier calcul. En ordonnant suivant $E_n^{(N)} \leq E_{n+1}^{(N)}$ les énergies obtenues avec N orbitales atomiques, un théorème dû à Fröbenius[35] permet d'affirmer que $E_n^{(N+1)} \leq E_n^{(N)}$; en particulier, les énergies $E_{n=1}^{(N)}$ de l'état fondamental forment une suite décroissante, toujours en accord avec le Principe variationnel.

On conçoit aisément que ces généralisations, qui améliorent sensiblement la description pour l'ion moléculaire considéré, s'étendent sans peine à des molécules plus complexes – sans préjuger par ailleurs du mode de traitement de l'interaction entre les électrons au sein de la molécule[36]. On en donne dans la suite trois exemples, choisis pour leur importance

[33] Avec $E_+ - E_- \sim 10$ eV, $\nu \sim 10^{15}$ Hz.

[34] Autre cas de stabilisation par résonance : la molécule de benzène, voir sous-section 28.2.4.

[35] aussi connu sous le nom de théorème de Rayleigh - Ritz.

[36] Pour une molécule contenant plusieurs électrons, on peut encore former des orbitales moléculaires comme ci-dessus, que l'on remplit les unes après les autres en épuisant le *stock* d'électrons et avec lesquelles on forme un déterminant de Slater. Il est clair que la fonction d'onde ainsi construite est caractéristique d'un modèle à électrons indépendants. On sait faire beaucoup mieux : de toute évidence, la base d'orbitales atomiques choisie judicieusement peut aussi être utilisée pour *représenter* (approximativement) les orbitales moléculaires de Hartree - Fock, afin de prendre (partiellement) en compte la répulsion entre les électrons, le schéma de Hartree-Fock n'étant qu'une approximation parmi d'autres.

historique, et pour les notions qu'ils permettent d'introduire, que l'on retrouvera plus tard à propos de la matière condensée.

28.2.3 La molécule d'hydrogène : la méhode de Heitler - London

Après H_2^+, l'édifice moléculaire le plus simple est la molécule d'hydrogène H_2, dont le premier traitement été accompli par Heitler et London en 1927 [221]. C'est aussi la première *molécule* où le Principe de Pauli peut jouer son rôle, dont l'importance magistrale a déjà pu être constatée à propos de l'hélium.

La question est ici de s'intéresser aux fonctions propres du Hamiltonien suivant :

$$\boxed{H \stackrel{\text{déf}}{=} H_{\text{ion}}(\vec{r}_1, \vec{p}_1) + H_{\text{ion}}(\vec{r}_2, \vec{p}_2) + \frac{e'^2}{r_{12}} + \frac{e'^2}{R}} \tag{28.67}$$

où $H_{\text{ion}}(\vec{r}, \vec{p})$ est le Hamiltonien électronique de l'ion H_2^+ et représente un électron tout seul dans le champ des deux protons, supposés fixes (première étape de Born et Oppeheimer), et situés à une distance R l'un de l'autre. Par rapport à H_2^+, une différence majeure tient au terme à deux corps, qui comme lorsque l'on passe de l'atome d'hydrogène à celui d'hélium, introduit une nouvelle difficulté d'envergure. C'est un fait entendu que, comme pour He et pour la même raison, on ne sait pas résoudre exactement l'équation aux fonctions propres $H\Psi = E\Psi$ avec H donné en (28.67). Le point est donc réellement de construire des approximations satisfaisantes.

S'agissant d'un système de deux électrons, on peut former des états triplet et singulet, dont la partie d'espace $\Psi(\vec{r}_1, \vec{r}_2)$, forcément factorisée de la partie de spin (voir chapitre 22, note 14 p. 998), est soit symétrique, soit antisymétrique :

$$\Psi_{\text{singulet}}(\vec{r}_1, \vec{r}_2) = +\Psi_{\text{singulet}}(\vec{r}_2, \vec{r}_1) \ , \quad \Psi_{\text{triplet}}(\vec{r}_1, \vec{r}_2) = -\Psi_{\text{triplet}}(\vec{r}_2, \vec{r}_1) \ , \tag{28.68}$$

Toute la question est de trouver une bonne fonction Ψ approchée.

La méthode de Heitler et London est dite *covalente* (*valence bond*) pour la raison suivante. Une première idée pour traiter le problème consisterait à repartir des deux orbitales moléculaires $\Phi_{E_\pm}$ obtenues pour H_2^+ et, s'agissant de décrire les premiers états de H_2, de former les bonnes combinaisons des produits de deux facteurs. Par exemple, pour décrire l'état fondamental, on pourrait choisir :

$$\Psi_{\text{fond}}(\vec{r}_1, \vec{r}_2) = \Phi_{E_+}(\vec{r}_1)\Phi_{E_+}(\vec{r}_2) \ , \tag{28.69}$$

et pour décrire deux états excités, l'un triplet, l'autre singulet :

$$\Psi_{S^*}(\vec{r}_1, \vec{r}_2) = C_\pm\big[\Phi_{E_+}(\vec{r}_1)\Phi_{E_-}(\vec{r}_2) \pm \Phi_{E_+}(\vec{r}_2)\Phi_{E_-}(\vec{r}_1)\big] \ , \tag{28.70}$$

avec le signe + pour le singulet $S = 0$, le signe − pour le triplet $S = 1$. Imaginons maintenant que l'on développe un produit des deux orbitales moléculaires $\Phi_{E_\varepsilon}(\vec{r}_1)\Phi_{E_{\varepsilon'}}(\vec{r}_2)$

figurant dans les formes (28.69) ou (28.70) ; compte tenu de leur expression (28.59), quatre termes apparaissent ; deux d'entre eux représentent schématiquement un électron au voisinage d'un proton, $\phi_1(\vec{r}_1)\phi_2(\vec{r}_2)$ et $\phi_1(\vec{r}_2)\phi_2(\vec{r}_1)$. Les deux autres termes, $\phi_i(\vec{r}_1)\phi_i(\vec{r}_2)$ avec $i = 1, 2$, correspondent à un proton *nu*, l'autre étant habillé avec deux électrons. Autrement dit, certains termes représentent deux atomes d'hydrogène côte-à-côte (termes dits *covalents*), les autres un ion H^- près d'un proton (termes *ioniques*). Selon (28.59), *tous* ces termes ont exactement le même poids, ce qui est visiblement très artificiel et ne correspond sûrement pas à une bonne description de la réalité physique : l'hydrogène est un édifice très stable[37], au contraire l'ion H^- est plutôt fragile (voir problème 27.8.2). Si l'on ne peut pas (doit pas !) exclure que des termes ioniques interviennent dans la *représentation* approchée d'une fonction d'onde moléculaire, on ne peut raisonnablement leur attribuer un rôle de même importance que celui dévolu aux termes covalents.

Le point de départ de Heitler et London est motivé par ces considérations physiques, et consiste à poser une fonction sous la forme :

$$\boxed{\Psi_{\text{HL}}(\vec{r}_1, \vec{r}_2) = C'_\pm\left[\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) \pm \phi_1(\vec{r}_2)\phi_2(\vec{r}_1)\right]} \tag{28.71}$$

où les ϕ_i sont cette fois des orbitales *atomiques*. Clairement, chaque terme est de nature *covalente*, par construction, et tout aspect *a priori* artificiel est exclu d'emblée. Avec $S \stackrel{\text{déf}}{=} \langle\phi_1|\phi_2\rangle$, la constante de normalisation $C'_\pm$ vaut $[2(1 \pm S^2)]^{-1/2}$.

Une toute première question, pour l'instant quelque peu secondaire, est le choix des orbitales atomiques ϕ_i. Il est très naturel d'utiliser l'état fondamental ψ_{100} de l'atome d'hydrogène, en centrant une fonction de ce type sur chaque proton. En désignant par $\vec{R}_i$ la position du proton P_i, on choisit donc les deux fonctions, supposées normalisées :

$$\phi_1(\vec{r}) \stackrel{\text{déf}}{=} \psi_{100}(\vec{r} - \vec{R}_1) \ , \qquad \phi_2(\vec{r}) \stackrel{\text{déf}}{=} \psi_{100}(\vec{r} - \vec{R}_2) \ , \tag{28.72}$$

se déduisant l'une de l'autre par la translation $\mathcal{T}(\vec{R}_1 - \vec{R}_2)$; le point de départ, au niveau *atomique*, est donc le même que pour H_2^+, l'image physique à l'ordre zéro étant celle de deux électrons chacun près de son proton. Lorsque la distance R entre ces derniers est très grande, l'édifice est constitué de deux objets en interaction attractive (forces de van der Waals), quoique l'énergie totale n'est pas minimum par rapport à R. Le rapprochement des deux atomes ainsi constitués va permettre à chaque électron d'aller visiter l'autre proton – par effet tunnel en quelque sorte – avec une probabilité qui croît très vite quand R diminue. Comme pour H_2^+, c'est cet échange d'électrons entre les protons qui est le fondement physique de la stabilité moléculaire ; les protons partageant *deux* électrons et non un seul, on doit s'attendre à obtenir une molécule plus stable que le fragile ion moléculaire H_2^+ – ce qui est bien sûr avéré expérimentalement[38].

Une fois ces éléments posés, la route est toute tracée : on se trouve dans un espace d'états effectifs de dimension égale à 2, engendré par le produit tensoriel des deux

---

[37] Par ailleurs, la liaison chimique de la molécule H_2 a toutes les caractéristiques d'une liaison covalente, ne serait-ce que parce que la symétrie interdit l'existence d'un moment dipolaire électrique.

[38] L'énergie de dissociation de H_2 est environ 109 kcal/mole, soit 4,71 eV – à comparer à 2,79 eV (voir (28.50)).

orbitales atomiques décalées $\phi_1(\vec{r})$ et $\phi_2(\vec{r})$:

$$\psi_1(\vec{r}_1, \vec{r}_2) \equiv \langle \vec{r}_1, \vec{r}_2 | \psi_1 \rangle = \phi_1(\vec{r}_1)\phi_2(\vec{r}_2) \ , \ \psi_2(\vec{r}_1, \vec{r}_2) \equiv \langle \vec{r}_1, \vec{r}_2 | \psi_2 \rangle = \phi_2(\vec{r}_1)\phi_1(\vec{r}_2) \ . \tag{28.73}$$

Il convient donc d'écrire la matrice de H donné par (28.67), et de la diagonaliser, soit de commencer par annuler le déterminant[39] $\mathrm{Det}|\langle\psi_i|H|\psi_j\rangle - ES_{ij}|$, où H est le Hamiltonien (28.67) et où $S_{ij} = S^2 + (1-S^2)\delta_{ij}$, S étant le produit scalaire $\langle\phi_1|\phi_2\rangle$, dont l'expression est donnée en (28.61). On a :

$$\langle\psi_1|H|\psi_1\rangle = \langle\psi_2|H|\psi_2\rangle = 2\tilde{\varepsilon} + \frac{e'^2}{R} + \langle\phi_1\phi_2|\frac{e'^2}{r_{12}}|\phi_1\phi_2\rangle \ , \tag{28.74}$$

et :

$$\langle\psi_1|H|\psi_2\rangle = \langle\psi_2|H|\psi_1\rangle = 2S\tilde{v} + S^2\frac{e'^2}{R} + \langle\phi_1\phi_2|\frac{e'^2}{r_{12}}|\phi_2\phi_1\rangle \ . \tag{28.75}$$

Les deux quantités $\tilde{\varepsilon}$ et $\tilde{v}$ sont :

$$\tilde{\varepsilon} \stackrel{\text{déf}}{=} \langle\phi_i|\frac{\vec{p}^2}{2m} - \frac{e'^2}{\|\vec{r} - \vec{R}_1\|} - \frac{e'^2}{\|\vec{r} - \vec{R}_2\|}|\phi_i\rangle \ , \tag{28.76}$$

$$\tilde{v} \stackrel{\text{déf}}{=} \langle\phi_1|\frac{\vec{p}^2}{2m} - \frac{e'^2}{\|\vec{r} - \vec{R}_1\|} - \frac{e'^2}{\|\vec{r} - \vec{R}_2\|}|\phi_2\rangle \ ; \tag{28.77}$$

En ce qui concerne les contributions des termes à un électron, elles ont toutes été déjà obtenues lors du calcul pour H_2^+. La seule nouveauté provient du terme à deux corps $\frac{e'^2}{r_{12}}$, qui force à calculer les deux intégrales suivantes, du type direct et d'échange comme pour l'atome d'hélium :

$$J \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} \mathrm{d}^3 r_1 \int_{\mathbb{R}^3} \mathrm{d}^3 r_2 \, [\phi_1(\vec{r}_1)]^2 \, \frac{e'^2}{\|\vec{r}_1 - \vec{r}_2\|} \, [\phi_1(\vec{r}_2)]^2 \ , \tag{28.78}$$

$$K \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} \mathrm{d}^3 r_1 \int_{\mathbb{R}^3} \mathrm{d}^3 r_2 \, \phi_1(\vec{r}_1)\phi_2(\vec{r}_2) \, \frac{e'^2}{\|\vec{r}_1 - \vec{r}_2\|} \, \phi_1(\vec{r}_2)\phi_2(\vec{r}_1) \ . \tag{28.79}$$

Avec ces notations, l'annulation du déterminant donnent les deux solutions pour l'énergie totale :

$$\boxed{E_{\mp}(R) = \frac{2(\tilde{\varepsilon} \pm S\tilde{v}) + J \pm K}{1 \pm S^2} + \frac{e'^2}{R}} \tag{28.80}$$

Le calcul de J et K s'effectue en développant d'abord $\frac{1}{r_{12}}$ en série de produits d'harmoniques sphériques (voir (18.123)), comme on l'a fait explicitement pour He (voir chapitre 27, section 27.6). En raison de la symétrie sphérique des fonctions ϕ_i, la série se réduit au seul terme $k = 0$; l'intégrale radiale restant à calculer s'obtient en passant en coordonnées elliptiques. On arrive ainsi sans trop de peine à l'expression de J :

$$J = \frac{e'^2}{R} - \frac{e'^2}{a_0}\left(\frac{1}{\rho} + \frac{11}{8} + \frac{3}{4}\rho + \frac{1}{6}\rho^2\right)\mathrm{e}^{-2\rho} \ , \quad \rho \stackrel{\text{déf}}{=} \frac{R}{a_0} \ . \tag{28.81}$$

[39] À nouveau, comme les $|\psi_i\rangle$ ne sont pas orthogonaux, la quantité $\langle\psi_i|H|\psi_j\rangle$ n'est pas un élément de la matrice de H sur la base $\{|\psi_i\rangle\}_i$.

Le calcul explicite de l'intégrale d'échange K est autrement plus ardu (voir [223], [215], p. 266), et produit une expression assez peu plaisante :

$$K = \frac{e'^2}{5a_0}\left\{\left(\frac{25}{8}-\frac{23}{4}\rho-3\rho^2-\frac{1}{3}\rho^3\right)\mathrm{e}^{-2\rho}+\frac{6}{\rho}\left[S^2(\mathrm{C}+\ln\rho)+S'^2\mathrm{Ei}(-4\rho)-2SS'\mathrm{Ei}(-2\rho)\right]\right\} . \tag{28.82}$$

Dans cette expression, la quantité S' est :

$$S' \stackrel{\text{déf}}{=} \left(1-\rho+\frac{1}{3}\rho^2\right)\mathrm{e}^{+\rho} . \tag{28.83}$$

Par ailleurs, C est la constante[40] d'Euler ($\mathrm{C} = 0,577...$), et $\mathrm{Ei}(-x)$ désigne l'*exponentielle intégrale*, fonction spéciale définie pour $x > 0$ comme ([38], § 8.211) :

$$\mathrm{Ei}(-x) \stackrel{\text{déf}}{=} -\int_x^{+\infty}\frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t . \tag{28.84}$$

Les deux intégrales J et K sont tracées pour la comparaison sur la figure 28.9. Tout naturellement, l'intégrale K tend vers zéro beaucoup plus vite que J puisqu'elle implique le produit $\phi_1\phi_2$ en interaction avec lui-même, ce produit allant vite (exponentiellement) à zéro aux grands R.

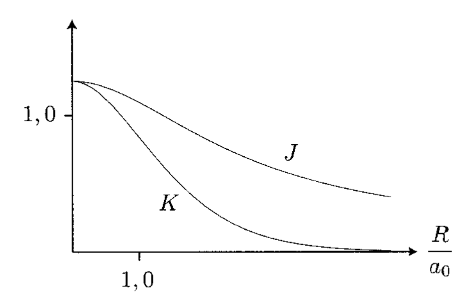


Figure 28.9: Variation des intégrales directe, J, et d'échange, K, en fonction de la distance R. L'unité en ordonnée est $\frac{e'^2}{2a_0}$.

Ce travail technique étant accompli, on connaît explicitement les quantités apparaissant dans (28.80), et donc les deux valeurs les propres $E_-(R) < E_+(R)$, quel que soit R ; les deux fonctions propres correspondantes $\Psi_{\mathrm{HL},\mp}$ ont forcément (par symétrie) l'allure donnée en (28.71) (pour plus de détails, voir problème 28.4.4). La fonction $\Psi_{\mathrm{HL},+}$, symétrique, correspond à un état singulet et est associée à l'énergie la plus basse $E_-(R)$; la fonction $\Psi_{\mathrm{HL},-}$ correspond à l'énergie la plus élevée[41] $E_+(R)$; étant antisymétrique (et visiblement antiliante), elle ne peut être utilisée que pour former un état triplet.

[40]On adopte toujours ici les notations de Gradsteyn et Rhyzhik [38], où $\mathrm{C} = 0,577\,215\,665...$ et où $\gamma \stackrel{\text{déf}}{=} \mathrm{e}^C = 1,781\,072\,418...$; dans certains ouvrages, les symboles C et γ sont échangés.

[41]Noter que, par comparaison avec un atome, la disposition singulet - triplet est inversée, ce qui permet de penser que la répulsion électronique est de moindre importance en raison de l'influence des *deux* protons séparés de la distance R. En outre, il est visible que le recouvrement $S(R)$ joue un rôle important en la matière, puisque l'intégrale d'échange (positive !) arrive avec le même signe que pour un atome, $-$ pour le triplet, $+$ pour le singulet.

Pour l'instant, aucune conclusion sur la *stabilité* moléculaire ne peut être tirée (existence d'un *minimum* en fonction de R ?), compte tenu de la complexité technique de l'expression des énergies $E_{\pm}(R)$. Il faut en fait en venir au tracé numérique des énergies, qui révèle deux courbes analogues à celles de la figure 28.8 (voir fig. 28.10), avec évidemment des valeurs numériques différentes. À ce stade, on peut donc conclure qualitativement que l'état fondamental (singulet) est stable, et que l'état excité triplet formé avec les mêmes orbitales atomiques est instable. Ce succès est malheureusement vite tempéré par le calcul des valeurs numériques correspondant au minimum de $E_-(R)$. On trouve que la distance d'équilibre vaut 0, 80 Å, alors qu'expérimentalement elle est égale à 0,75 Å ; quant à l'énergie de dissociation, elle est fort décevante : 73 kcal/mole (environ 3,2 eV), au lieu de 109 kcal/mole (4,72 eV)... L'erreur est donc considérable : la méthode produit une molécule un peu trop grosse, et surtout bien moins robuste[42] qu'elle n'est en réalité. Bien évidemment, la fonction $E_-(R)$ est au-dessus de la valeur exacte, ce que confirment des mesures avec R, pour autant que l'on puisse faire varier R en pratique. Au total, la méthode première de Heitler et London, séduisante sur le plan des principes et fortement inspirée par des considérations physiques légitimes, se révèle loin d'être satisfaisante quantitativement.

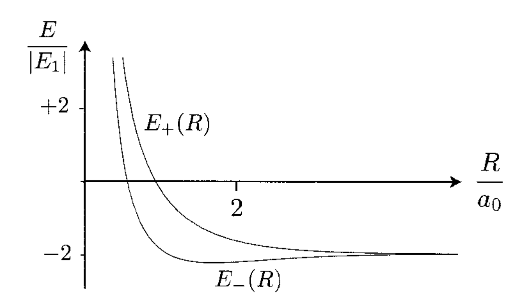


Figure 28.10: Énergies $E_{\pm}(R)$ de la théorie de Heitler et London.

Cet échec relatif pose la question d'une description plus élaborée. Ceci peut se faire de mille façons – on aurait bien du mal à compter les nombreux articles publiés sur ce sujet (pour un article récent, voir [222]) –, mais l'une d'entre elles s'impose à l'esprit, pas seulement pour son *apparente* simplicité, compte tenu de l'expérience acquise avec l'hélium (voir chapitre 27, section 27.6) : mettre un facteur d'échelle Z dans les orbitales atomiques, et le faire varier en s'appuyant sur le Principe variationnel. On sait qu'alors, non seulement la valeur de l'énergie sera forcément plus proche de la valeur exacte, mais en plus la fonction ainsi obtenue satisfera, par construction, le théorème du Viriel – ce qui, comme on l'a vu, n'est toutefois pas un gage de qualité (chapitre 27, section (27.5)). Wang [223] a effectué ce calcul pour la distance d'équilibre, et l'amélioration est très importante. Peu après, Rosen [224] a généralisé ceci à une distance internucléaire R quelconque, trouvant que, en fonction de R, le paramètre $Z_{\text{opt}}(R)$ a une décroissance d'allure exponentielle, partant de la valeur[43] $Z_{\text{opt}}(0) = \frac{27}{16}$, et décroissant à $Z_{\text{opt}} \simeq 1$

[42]En raison du Principe variationnel, l'énergie de liaison est forcément toujours plus petite que sa valeur exacte.

[43]En $R = 0$, on a un atome d'hélium ; la variation du facteur d'échelle reproduit naturellement ce que

pour $R \sim a_0$. Muni de ce résultat, le point important est d'examiner la variation avec R de la fonction $E_+(R,\, Z_{\mathrm{opt}}(R))$. On trouve[44] qu'elle est minimum pour $R \sim 0,74$ Å, valeur très proche de la valeur expérimentale, d'où l'on tire une énergie de dissociation voisine de 3,78 eV, en nette amélioration comparativement au traitement de Heitler - London, même s'il manque encore près d'un eV.

Il n'est pas aisé de cerner les causes de ce (relatif) insuccès, et ce d'autant plus que la formation de la liaison chimique s'accompagne de compensations subtiles entre les différentes contributions à l'énergie totale, comme il sera montré dans la sous-section 28.2.6 ; une analyse qualitative purement statique (ne raisonnant qu'avec l'énergie potentielle) est donc sans objet. Une amélioration certaine est à attendre dans la *forme* des orbitales atomiques elles-mêmes : le choix de fonctions sphériques, s'il ne viole pas la symétrie du problème, ne prend pas en compte le fait que, pour chaque proton, la présence de l'autre déforme le nuage électronique. Une façon d'autoriser une telle polarisation est d'inclure des harmoniques sphériques $l = 1$ judicieusement choisies, autrement dit à ajouter à la base d'états des fonctions du genre $R_{21}Y_{1m}$, donc de type 2p. Une autre possibilité est de réintroduire les configurations ioniques délaissées par Heitler et London, mais en laissant *flotter* leur poids relativement aux termes covalents.

Une autre cause certaine de l'insuccès est évidemment le traitement très rustique de la corrélation électronique puisque, tout compte fait, on s'est borné à inclure l'effet de $\frac{e'^2}{r_{12}}$ par une perturbation au premier ordre. Pour l'hélium, on a constaté que cette approximation était pour le moins insuffisante, et il en est évidement de même ici. L'amélioration immédiate consiste à laisser libres au départ les orbitales atomiques, et à les déterminer par une méthode auto-cohérente du genre SCF.

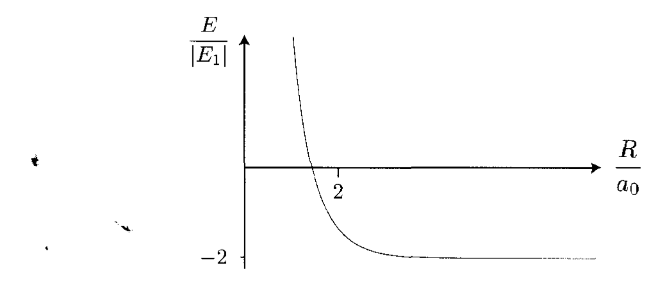


Figure 28.11: Variation de l'énergie totale en l'inexistence du Principe de Pauli. L'énergie n'a pas de minimum, la molécule d'hydrogène H_2... n'existe pas !

Quoi qu'il en soit des schémas très variés permettant d'obtenir une description plus satisfaisante, une chose est sûre : le Principe de Pauli, plus précisément le postulat d'antisymérisation, joue à nouveau un rôle crucial. En l'absence de ce postulat – qui exige que la fonction (d'espace) du fondamental est symétrique –, il n'y aurait aucune raison de former la combinaison linéaire + figurant dans (28.71) ; les considérations physiques

l'on a alors trouvé pour cet atome dans le même cadre d'approximation (voir (27.46)).

[44] On imagine la patience qu'il a fallu pour en arriver là, à une époque où les moyens de calcul numérique étaient quasi inexistants.

de Heitler et London conduiraient alors à retenir le simple produit $\phi_1(\vec{r}_1)\phi_2(\vec{r}_2)$. Avec ce choix, il n'y a qu'une seule fonction disponible et le seul travail à effectuer, gardant par ailleurs exactement le même cadre approximatif, est de calculer la valeur moyenne du Hamiltonien avec cette seule et unique fonction. En réalité, ce travail est déjà fait : il suffit de faire disparaître partout l'intégrale d'échange qui, avec un simple produit, ne peut, de toute évidence, pas survenir. Le résultat est spectaculaire : la seule et unique énergie totale, $E(R)$, n'a plus de minimum (voir fig. 28.11), avec pour conclusion que la molécule H_2 n'existe pas. Le rôle du Principe de Pauli, et donc de cette intégrale d'échange K si ardue à calculer, se révèle une fois encore de la plus haute importance.

28.2.4 La molécule de benzène

De toute évidence, la méthode de Heitler et London ne se généralise pas simplement pour une molécule plus complexe. On va voir ici, en prenant l'exemple du benzène, comment il est possible de traiter de façon élémentaire une molécule nettement plus grosse.

Expérimentalement, on sait que le benzène, de formule C_6H_6, a une géométrie hexagonale. Les 6 atomes de carbone sont aux sommets d'un hexagone régulier ; la distance entre deux atomes voisins est voisine de $1,4$ Å, intermédiaire entre une liaison simple ($1,45$ Å) et une liaison double ($1,34$ Å). Les liaisons partant d'un même carbone forment donc un angle de 120° ; en outre, chacun est lié à un atome d'hydrogène, la liaison C–H étant également orientée suivant le même angle par rapport aux liaisons adjacentes, et a pour longueur $1,1$ Å (voir fig. 28.12).

Le carbone ($Z = 6$) possède 6 électrons et a pour configuration électronique $(1s)^2(2s)^2(2p)^2$. Les deux électrons (1s) sont liés au noyau avec une énergie de l'ordre[45] de $(6)^2|E_1|$, soit près de 500 eV ; le premier niveau excité du carbone (estimé toujours avec les niveaux hydrogénoïdes $Z = 6$) est à environ 370 eV au-dessus du fondamental. Compte tenu de l'ordre de grandeur des processus physico-chimiques (~ 1 eV), les électrons (1s) ne peuvent pas être excités et sont, pour ces mécanismes, quasi-inertes. Avec le noyau, il forment ce *cœur* qui a été qualitativement défini au début de ce chapitre. Restent ainsi les quatre électrons de la couche L ($n = 2$), qui ont tous la même énergie de liaison, situés dans une couche très malléable puisqu'elle est incomplète. Comment s'organisent-ils ?

On ne peut répondre à cette question sans arguments de nature empirique, le carbone étant connu pour se lier majoritairement soit de façon tétraédrique (dans la forme du diamant à l'état solide, ou du méthane en phase vapeur), soit en formant des angles de 120° comme dans le benzène justement. On explique ces faits, qui montrent bien que la liaison chimique a un caractère *directionnel* très marqué[46], en construisant des combinaisons linéaires des orbitales de la couche 2p ayant précisément cette géométrie dirigée. On dispose de quatre fonctions qui, gardant la notation hydrogénoïde, sont

[45]Pour un atome hydrogénoïde Z, l'énergie est $E_n(Z) = Z^2 E_n$, avec $E_n \simeq -13,6/n^2$ eV.

[46]Cette caractéristique directionnelle est responsable pour une grande part de la *rigidité* d'une molécule.

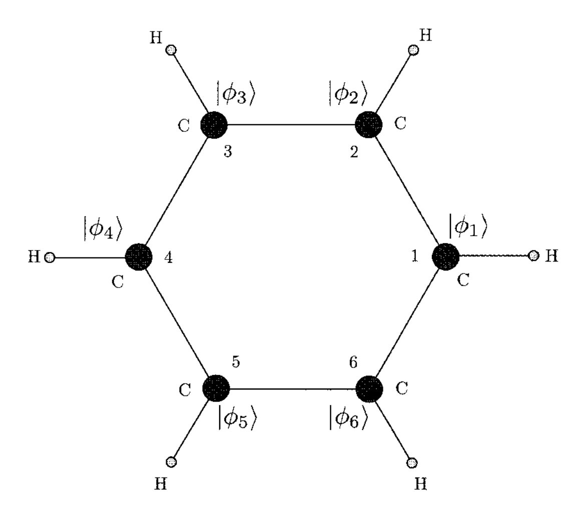


Figure 28.12: La molécule de benzène.

$R_{20}Y_{00}$ et $R_{21}Y_{1m}$; Y_{10}, étant proportionnelle à $\cos\theta$, présente des lobes de densité maximale le long de l'axe de quantification ($\theta = 0,\ \pi$). Comme les Y_{1m} se transforment les unes dans les autres par rotation, on peut aussi construire, par combinaison linéaire, deux autres fonctions ayant des maxima de densité dans le plan perpendiculaire et même faire en sorte qu'elles soient orientées à 120° l'une de l'autre. Au total, partant des quatre fonctions ci-dessus, on peut définir, pour un atome de carbone, une orbitale atomique $|\phi_z\rangle$ pointant principalement dans la direction Oz et trois autres $|\chi_i\rangle$ ($i = 1,\ 2,\ 3$) situées dans le plan perpendiculaire et formant une étoile régulière à trois branches. Ces orbitales sont dites *hybrides* et leur construction est une opération appelée *hybridation des orbitales*. L'hybridation particulière[47] conduisant à l'étoile à trois branches est dite *hybridation* sp^2. De toute évidence, elle est appropriée à la géométrie du benzène.

Admettant ceci, la formation de la molécule peut se décrire de façon imagée comme suit, définissant Oz comme l'axe perpendiculaire à l'hexagone. Chaque carbone entrant dans le cycle benzénique arrive avec ses trois orbitales hybrides abritant chacune un électron ; chacun de ces trois électrons va pouvoir s'apparier d'une part avec les deux autres venant des carbones voisins, d'autre part avec celui associé à l'atome d'hydrogène. Ce jeu de construction où tout s'emboîte bien constitue le squelette benzénique, et il reste au total les 6 électrons dans l'orbitale $|\phi_z\rangle$ de chaque carbone, tous équivalents et relativement mobiles par rapport aux autres, lesquels sont désormais engagés dans les liaisons chimiques dont l'énergie de rupture est typiquement de quelques eV (quelques centaines de kJ/mole).

Une fois la molécule constituée, un premier modèle pour la décrire consiste donc à ne retenir que les 6 électrons non appariés, tous les autres étant très solidement liés – dans le cœur ou dans les liaisons C – C et C – H. La nature très fortement empirique du

[47] Pour le diamant (ou le méthane), on combine linéairement les quatre orbitales de la couche L, pour former une étoile joignant le centre d'un tétraèdre régulier à ses quatre sommets : c'est l'hybridation sp^3.

cadre ainsi défini permet néanmoins de dégrossir la physique du benzène et, finalement, d'expliquer bien des choses, sans compter avec les notions qui vont s'introduire d'elles-mêmes et que l'on retrouvera plus tard dans un contexte différent (polyacétylène, sous-section 28.2.5), puis théorème de Bloch (chapitre 30, section 30.2). Bien sûr, les autres électrons ne sont pas totalement hors-jeu : rhabillant les interactions, ils contribuent à assurer l'existence d'une configuration moléculaire stable et relativement rigide, dont toute déformation coûte de l'énergie.

Ceci étant admis, on a un système à six électrons se promenant dans un potentiel possédant la symétrie hexagonale. Le Hamiltonien correspondant commute donc avec toutes les opérations (géométriques) de symétrie de l'hexagone, en particulier avec les rotations de $\frac{2\pi}{6} = \frac{\pi}{3}$ autour d'un axe passant par le centre de l'hexagone et perpendiculaire au plan de celui-ci. L'espace des états de chacun de ces six électrons est *a priori* de dimension infinie, puisque sur chaque atome on peut introduire autant d'orbitales atomiques que l'on veut. Compte tenu de la rusticité du modèle déjà bien cadré, on se borne à introduire sur chaque carbone la seule et unique orbitale de type $|\phi_z\rangle$ issue de la couche L. Pour la simplicité des notations, on note $|\phi_n\rangle$ l'orbitale centrée sur le carbone n (voir fig 28.12). On définit ainsi un espace d'états effectif, $\mathcal{E}$, de dimension égale à 6.

On va maintenant exploiter à fond la symétrie du problème. Soit R la rotation dans $\mathbb{R}^3$ de $+\frac{\pi}{3}$; à cette rotation dans l'espace physique correspond une transformation unitaire $\mathcal{R}$ dans l'espace des états $\mathcal{E}$ suivant la prescription habituelle (voir (17.93)) :

$$\phi'(\vec{r}) \overset{\text{déf}}{=} \mathcal{R}\,\phi(\vec{r}) = \phi(R^{-1}\,\vec{r}) \ . \tag{28.85}$$

Avec cette définition et la numérotation de la figure (17.93), on a :

$$\mathcal{R}\,\phi_n(\vec{r}) = \phi_{n+1}(\vec{r}) \ , \tag{28.86}$$

ce que l'on peut aussi noter $\mathcal{R}|\phi_n\rangle = |\phi_{n+1}\rangle$, et plus généralement :

$$\boxed{\mathcal{R}^m|\phi_n\rangle = |\phi_{n+m}\rangle} \tag{28.87}$$

Le Hamiltonien H commute avec les 6 rotations distinctes $\mathcal{R}^m$ ($m = 1, 2, ..., 6$), qui forment un groupe abélien ; si on arrive à former d'emblée les vecteurs propres de la rotation $\mathcal{R}$, la matrice de H sur cette base sera prédiagonalisée[48]. Ainsi, et quels que soient les détails définissant H ultérieurement, on aura résolu le problème de la recherche des vecteurs propres par le seul usage de la symétrie.

Soit donc à trouver les vecteurs $|\psi\rangle$ propres de $\mathcal{R}$, exprimés sur la base des $\{|\phi_n\rangle\}_n$:

$$|\psi\rangle = \sum_{n=1}^{6} c_n|\phi_n\rangle \ , \tag{28.88}$$

[48]En fait, on la trouvera *complètement* diagonale, en conséquence du fait que le groupe de symétrie est *abélien*, et que l'espace d'états effectif est justement de dimension égale à 6.

et satisfaisant $\mathcal{R}|\psi\rangle = r|\psi\rangle$, $r \in \mathbb{C}$; les $\psi(\vec{r})$ seront les orbitales moléculaires construites sur les orbitales atomiques $\phi_n(\vec{r})$ – on est donc typiquement dans un schéma LCAO. On sait d'avance que la valeur propre r est un complexe de module 1, puisque $\mathcal{R}$ est unitaire ; ceci permet de poser d'emblée $r \stackrel{\text{déf}}{=} \mathrm{e}^{-\mathrm{i}\alpha}$ avec $\alpha \in \mathbb{R}$. En appliquant maintenant $\mathcal{R}$ aux deux membres, on a :

$$\mathcal{R}|\psi\rangle = \sum_{n=1}^{6} c_n \mathcal{R}|\phi_n\rangle = \sum_{n=1}^{6} c_n|\phi_{n+1}\rangle \ , \tag{28.89}$$

où l'égalité (28.86) été utilisée. En décalant les indices, on voit que la somme de droite est égale à[49] $\sum_{n=2}^{7} c_{n-1}|\phi_n\rangle$; quant au premier membre, on veut qu'il soit $r|\psi\rangle$, soit $\sum_n rc_n|\phi_n\rangle$. Identifiant les coefficients, on en déduit $rc_n = c_{n-1}$, soit $c_n = \mathrm{e}^{\mathrm{i}\alpha}c_{n-1}$. De proche en proche, prenant c_1 comme référence (c'est comme une *condition initiale* pour la récurrence entre les coefficients), on a $c_n = \mathrm{e}^{\mathrm{i}(n-1)\alpha}c_1$. Prenant $n = 7$, il vient $c_7 = \mathrm{e}^{6\mathrm{i}\alpha}c_1$, mais $c_7 \equiv c_1$, d'où :

$$\boxed{\mathrm{e}^{6\mathrm{i}\alpha} = 1 \quad \Longleftrightarrow \quad \alpha = (\text{entier} \in \mathbb{Z}) \times \frac{2\pi}{6}} \tag{28.90}$$

Afin d'alléger les écritures, on pose $\theta \stackrel{\text{déf}}{=} \frac{2\pi}{6}$ pour la suite ; notant k l'entier de (28.90), on a ainsi obtenu les vecteurs propres de $\mathcal{R}$, que l'on peut justement indexer avec k :

$$|\psi_k\rangle = C\sum_{n=1}^{6} \mathrm{e}^{\mathrm{i}kn\theta}|\phi_n\rangle \ , \quad \mathcal{R}|\psi_k\rangle = \mathrm{e}^{\mathrm{i}k\theta}|\psi_k\rangle \ . \tag{28.91}$$

Il n'y a visiblement que 6 vecteurs $|\psi_k\rangle$ distincts (θ est défini à 2π près), que l'on a tous sans répétition en prenant (par exemple) $k = 0, 1, 2, ..., 5$. Deux quelconques de ces vecteurs correspondent à deux valeurs propres *distinctes* de $\mathcal{R}$; comme $\mathcal{R}$ et H commutent, on en déduit :

$$\boxed{\langle\psi_k|H|\psi_{k'}\rangle = 0 \quad \forall\, k \neq k'} \tag{28.92}$$

Ainsi, en jouant uniquement avec la symétrie, on a trouvé une base sur laquelle la matrice de H est diagonale : il ne reste plus qu'à calculer ces éléments diagonaux pour avoir les valeurs propres de H, c'est-à-dire les énergies du modèle. Cet acquis n'est en aucune façon tributaire des détails concernant H, qu'il faudra bien préciser tôt ou tard dans la suite, puisque seule la *symétrie* (hexagonale) du problème a été utilisée.

Avant de faire ce calcul, montrons que les vecteurs propres ainsi trouvés peuvent être mis sous une forme remarquable, anticipant le théorème de Bloch. Il est toujours loisible d'écrire $\psi_k(\vec{r})$ sous la forme :

$$\psi_k(\vec{r}) \stackrel{\text{déf}}{=} \mathrm{e}^{\mathrm{i}\theta f(\vec{r})}u_k(\vec{r}) \ , \tag{28.93}$$

où $f(\vec{r})$ et $u_k(\vec{r})$ sont des fonctions pour l'instant quelconques. Ayant posé la définition (28.93), on a aussi :

$$\psi_k(R^{-1}\vec{r}) = \mathrm{e}^{\mathrm{i}\theta f(R^{-1}\vec{r})}u_k(R^{-1}\vec{r}) \ ; \tag{28.94}$$

[49] En raison de la *cyclicité* du problème, $|\phi_{6+n}\rangle \equiv |\phi_n\rangle$.

le premier membre est $\mathcal{R}\psi_k(\vec{r}) = \mathrm{e}^{\mathrm{i}k\theta}\psi_k(\vec{r})$ en raison de (28.91), soit $\mathrm{e}^{\mathrm{i}k\theta}\mathrm{e}^{\mathrm{i}\theta f(\vec{r})}u_k(\vec{r})$ d'après la définition (28.93). Il vient ainsi :

$$\mathrm{e}^{\mathrm{i}k\theta}\mathrm{e}^{\mathrm{i}\theta f(\vec{r})}u_k(\vec{r}) = \mathrm{e}^{\mathrm{i}\theta f(R^{-1}\vec{r})}u_k(R^{-1}\vec{r}) \Longleftrightarrow u_k(R^{-1}\vec{r}) = \mathrm{e}^{\mathrm{i}\theta[k+f(\vec{r})-f(R^{-1}\vec{r})]}\,u_k(\vec{r}) \ ; \tag{28.95}$$

Choisissons maintenant la fonction $f(\vec{r})$ de sorte que $f(R^{-1}\vec{r}) = k + f(\vec{r})$; cela étant fait, on obtient $u_k(R^{-1}\vec{r}) = u_k(\vec{r})$, soit $\mathcal{R}u_k(\vec{r}) = u_k(\vec{r})$: la fonction $u_k(\vec{r})$ ainsi définie a ceci de remarquable qu'elle est *invariante* par les rotations discrètes de l'hexagone. C'est la préfiguration du théorème de Bloch, que l'on retrouvera à propos de la symétrie de translation, également discrète, d'un réseau (voir chapitre 30, section 30.2).

On note aussi que les $|\psi_k\rangle$ peuvent finalement s'écrire, par exemple :

$$|\psi_k\rangle = C\Big(\sum_{n=1}^{6}\mathrm{e}^{\mathrm{i}kn\theta}\mathcal{R}^{n-1}\Big)|\phi_1\rangle \ , \tag{28.96}$$

expression montrant que chaque $|\psi_k\rangle$, vecteur ayant la bonne symétrie, peut s'obtenir en appliquant une certaine combinaison linéaire des opérateurs du groupe à un vecteur $|\phi_1\rangle$ quelconque vis-à-vis de cette symétrie : cette procédure est exactement de même nature que celle utilisée dans le chapitre 22 pour engendrer une fonction ayant la symétrie voulue en lui appliquant les projecteurs Y_{I} (voir (22.44)).

Retournant à la question des valeurs propres, et en vertu de (28.92), les énergies propres ε_k sont égales à $\langle\psi_k|H|\psi_k\rangle$:

$$\varepsilon_k = \langle\psi_k|H|\psi_k\rangle = |C|^2\sum_{n,n'=1}^{6}\mathrm{e}^{\mathrm{i}(n'-n)k\theta}\langle\phi_n|H|\phi_{n'}\rangle \ , \tag{28.97}$$

la constante de normalisation étant telle que :

$$|C|^2\sum_{n,n'=1}^{6}\mathrm{e}^{\mathrm{i}(n'-n)k\theta}\langle\phi_n|\phi_{n'}\rangle = 1 \ . \tag{28.98}$$

Compte tenu de (28.97) et (28.98), les valeurs propres ont alors pour expression :

$$\boxed{\varepsilon_k = \frac{\sum_{n,n'=1}^{6}\mathrm{e}^{\mathrm{i}(n'-n)k\theta}\langle\phi_n|H|\phi_{n'}\rangle}{\sum_{n,n'=1}^{6}\mathrm{e}^{\mathrm{i}(n'-n)k\theta}\langle\phi_n|\phi_{n'}\rangle}} \tag{28.99}$$

À ce stade, et pour voir à quoi ressemblent les ε_k, force est de rentrer dans les détails du modèle, c'est-à-dire de préciser ce que valent les éléments de matrice $\langle\phi_{n'}|H|\phi_n\rangle$, ainsi que les produits scalaires $\langle\phi_n|\phi_{n'}\rangle$. Posons :

$$\langle\phi_n|H|\phi_n\rangle \stackrel{\text{déf}}{=} E_0 \ , \quad \langle\phi_n|H|\phi_{n+q}\rangle \stackrel{\text{déf}}{=} v_q \ , \quad \langle\phi_n|\phi_{n+q}\rangle \stackrel{\text{déf}}{=} S_q \quad (q=1,2,3) \ , \tag{28.100}$$

où bien sûr, ni v_q ni E_0 ne dépendent de n, label de l'atome considéré (tous les sites sont équivalents) ; comme les $\phi_n(\vec{r})$ sont des fonctions réelles, toutes les quantités introduites le sont aussi. L'expression (28.99) s'écrit alors[50] :

$$\boxed{\varepsilon_k = E_0 + \frac{2\sum_{q=1}^{3} v_q \cos kq\frac{\pi}{3}}{1 + 2\sum_{q=1}^{3} S_q \cos kq\frac{\pi}{3}} \qquad (k = 0,\, 1,\, ...,\, 5)} \tag{28.102}$$

Toutes les quantités avec $n \neq n'$ impliquent des fonctions atomiques localisées sur des sites *distincts*, et décroissent vite avec la distance entre deux sites : $|v_3| \ll |v_2| \ll |v_1|$. C'est pourquoi on peut écrire une expression approchée en ne retenant que v_1, et en négligeant, au dénominateur, les produits scalaires S_q devant 1 :

$$\varepsilon_k \simeq E_0 + 2v_1 \cos k\frac{\pi}{3} \qquad (k = 0,\, 1,\, ...,\, 5) \tag{28.103}$$

Avec cette simplification, on obtient les 6 orbitales moléculaires :

$$|\psi_k\rangle = \frac{1}{\sqrt{6}} \sum_{n=1}^{6} \mathrm{e}^{ikn\frac{\pi}{3}} |\phi_n\rangle \ , \qquad \mathcal{R}|\psi_k\rangle = \mathrm{e}^{ik\theta}|\psi_k\rangle \ , \tag{28.104}$$

et les 4 valeurs distinctes de l'énergie :

$$\varepsilon_0 \simeq E_0 + 2v_1 \ , \quad \varepsilon_1 = \varepsilon_5 \simeq E_0 + v_1 \ , \quad \varepsilon_2 = \varepsilon_4 \simeq E_0 - v_1 \ , \quad \varepsilon_3 \simeq E_0 - 2v_1 \ , \tag{28.105}$$

réparties dans l'intervalle $[E_0 - 2|v_1|,\, E_0 + 2|v_1|]$ (qui préfigure une *bande d'énergie*). Dans ces orbitales moléculaires, il faut caser 6 électrons ; compte tenu du spin, trois d'entre elles suffisent et, s'agissant de construire l'état fondamental de la molécule de benzène, il faut prendre les trois orbitales de plus basse énergie. En supposant[51] $v_1 < 0$, l'énergie des six électrons, délaissant leur répulsion mutuelle, est égale à $2(\varepsilon_0 + \varepsilon_1 + \varepsilon_5)$, d'où :

$$\boxed{E_{\text{fond}} = 6E_0 - 8|v_1|} \tag{28.106}$$

En définitive, et de façon imagée, l'élément de matrice v_1 est l'amplitude permettant à l'électron du site n de sauter vers les premiers voisins $n \pm 1$; en quelque sorte, c'est ce qui lui donne de l'énergie de nature cinétique. En même temps, c'est cette liberté de mouvement qui conduit à un abaissement de l'énergie des 6 électrons égal à $8|v_1|$. Quand les électrons ont la possibilité de se déplacer, au lieu de rester chacun sur son carbone, l'énergie diminue, paradoxalement[52] ; cette diminution de l'énergie est parfois appelée

[50] Noter que l'expression (28.102) se généralise immédiatement à une molécule à N atomes disposés sur un polygone régulier ; avec N pair, on a :

$$\varepsilon_k = E_0 + \frac{2\sum_{q=1}^{N/2} v_q \cos kq\frac{2\pi}{N}}{1 + 2\sum_{q=1}^{N/2} S_q \cos kq\frac{2\pi}{N}} \qquad (k = 0,\, 1,\, ...,\, N-1) \tag{28.101}$$

[51] Ceci correspond à la convention habituelle, à laquelle on peut toujours revenir en rephasant $|\phi_n\rangle$ en $(-1)^n|\phi_n\rangle$.

[52] Dans la sous-section 28.2.6, on discute l'évolution des valeurs moyennes des énergies cinétique et potentielle lors de la formation d'une liaison chimique.

stabilisation par résonance – les 6 vecteurs $|\phi_n\rangle$ ont tous la même énergie E_0, ils sont[53] *résonnants.* La description très simplifiée qui vient d'être exposée est traditionnellement appelée *méthode de Hückel* [225].

Les orbitales moléculaires $\psi_k(\vec{r})$ sont assez particulières, au sens où toutes donnent une densité de probabilité *diffuse*, répartie uniformément sur tout le cycle benzénique – dans chaque $|\psi_k\rangle$, chaque $|\phi_n\rangle$ a le même poids –, au lieu d'être concentrée entre deux atomes, comme dans une liaison chimique ordinaire ; cette particularité est à l'origine des propriétés dites *aromatiques* du benzène (et de ses molécules cousines), étant bien entendu que la description fine de ces propriétés exige des traitements autrement plus élaborés.

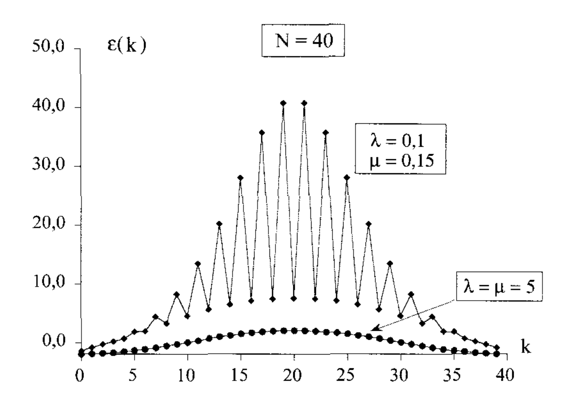


Figure 28.13: Variabilité de la loi de dispersion des valeurs propres selon la décroissance lente (♦) ou rapide (•) des $\langle\phi_n|H|\phi_{n+q}\rangle$ et des $\langle\phi_n|\phi_{n+q}\rangle$ avec la distance q.

Noter que si les orbitales atomiques ne sont pas assez concentrées (ramassées sur elles-mêmes), il faut garder l'expression complète (28.101), ou (28.102) ; il peut alors en résulter des modifications importantes de la dispersion des valeurs propres en fonction de l'entier k, puisqu'alors les numérateur et dénominateur de ces expressions peuvent varier assez vite en fonction de k. Afin de préciser ce point, admettons que toutes les quantités décroissent exponentiellement avec la distance, ce qui est raisonnable compte tenu du caractère localisé des orbitales atomiques :

$$v_q = v_1 \mathrm{e}^{-\lambda(q-1)} \quad , \quad S_q = \mathrm{e}^{-\mu q} \qquad (1 \leq q \leq \mathrm{E}[N/2]) \ , \tag{28.107}$$

λ et μ qualifiant précisément la portée de ces quantités par rapport à la distance entre sites. Si λ, $\mu \gg 1$, la formule approchée $\varepsilon_k \simeq 2|v_1| \cos k\frac{2\pi}{N}$ est tout à fait acceptable ; le cas intéressant est celui où ces deux nombres sont $\lesssim 1$. Avec le choix-modèle (28.107), il est facile de calculer les sommes figurant dans (28.101) pour une molécule

[53]La résonance peut aussi se voir entre les deux configurations de l'hexagone, miroirs l'une de l'autre, où on a dessiné une alternance de liaisons simples et doubles.

cyclique à N atomes (noter que si $N \gg 1$, l'aspect discret de l'entier k s'efface quelque peu). Au total, la loi de dispersion peut être profondément modifiée (voir fig. 28.13) : la courbe presque plate correspond à des amortissements très rapides, au contraire, la loi de dispersion oscillante est le résultat de la décroissance lente dans l'espace des quantités non-diagonales. Cette dualité de comportement reflète la portée spatiale des grandeurs impliquées, tout comme pour une transformée de Fourier : courte portée/spectre plat, longue portée/spectre structuré.

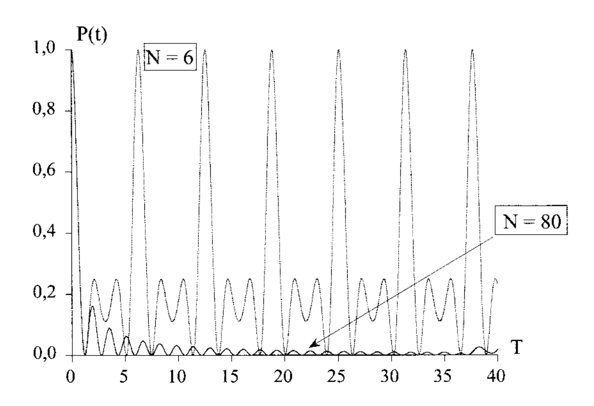


Figure 28.14: Probabilité en fonction du temps de retrouver un électron dans son orbitale atomique localisée de départ, selon la valeur de N (voir le texte) ; T est le temps réduit $\hbar^{-1}|v_1|t$. La courbe est tracée avec l'expression approchée $\varepsilon_k = 2|v_1|\cos k\frac{2\pi}{N}$.

Une particularité importante du spectre de H est d'être constitué, en général, d'énergies – donc de fréquences – non commensurables, et fournit un exemple simple d'évolution dynamique très variable selon la répartition précise des valeurs propres. Pour cela, imaginons qu'il est possible de préparer un électron dans l'une des orbitales atomiques localisées. Cet état de départ n'étant pas stationnaire, il se développe non trivialement au cours du temps ; l'état à l'instant t est :

$$|\Psi(t)\rangle = C\sum_{k=1}^{N} \mathrm{e}^{-\mathrm{i}\omega_k t}|\psi_k\rangle\langle\psi_k|\Psi(0)\rangle \qquad (\omega_k \stackrel{\text{déf}}{=} \hbar^{-1}\varepsilon_k) \ . \tag{28.108}$$

On en déduit la probabilité à l'instant t, $P(t)$, de trouver l'électron sur son site de départ :

$$P(t) \stackrel{\text{déf}}{=} |\langle\Psi(0)|\Psi(t)|^2 = \left|C\sum_{k=1}^{N} \mathrm{e}^{-\mathrm{i}\omega_k t}\right|^2 \ ; \tag{28.109}$$

quand N est petit, les fréquences sont commensurables, et le mouvement est périodique (voir le cas $N = 6$ sur la fig. 28.14). Au contraire, quand N est grand, apparaissent

forcément des fréquences *irrationnelles*[54] dans le spectre, donnant une évolution pseudo-périodique, avec des récurrences d'autant plus faibles et d'autant plus tardives que N est grand (voir fig. 28.14). C'est un exemple montrant comment une pseudo-irréversibilité peut survenir avec un nombre *fini* de fréquences[55].

28.2.5 Le polyacétylène

Le polyacétylène est un polymère de l'acétylène C_2H_2 ; on en connaît au moins deux formes, la forme dite *trans* (fig. 28.15) et la forme *cis*. On ne parlera ici que de la première en raison de l'intérêt qu'elle présente. En effet, on a découvert[56] en 1977 [226] que ce *plastique*, une fois dopé, pouvait devenir... conducteur, voyant sa conductivité augmenter d'un facteur 10^9 pour devenir comparable à celle du cuivre. Outre les questions fondamentales que pose ce matériau – dont la physique est fort complexe –, son intérêt pratique est évident : depuis l'invention de la batterie au plomb, on rêve de batteries *légères*.

Les phénomènes à l'origine de cette propriété spectaculaire sont en effet fort complexes ; une caractéristique très importante est l'existence de deux structures *dégénérées* : il est bien évident que la structure représentée sur la figure 28.15, et celle qui s'en déduit par une symétrie miroir dans un plan vertical, ont la même énergie. Dans le contexte de ce chapitre, il sera juste question ici de donner quelques idées sur la nature du matériau *pur*, qui apparaîtront naturellement comme une sorte de généralisation de ce qui vient d'être vu à propos du benzène, avec toutefois des caractéristiques propres, et qui décrivent un peu en détail le polyacétylène, devenu le prototype des *polymères conducteurs*. L'un des modèles favoris pour le décrire est dû à Schrieffer et al. [228], où toute corrélation électronique est ignorée malgré son importance [229], et les possibilités qu'elle ouvre (voir problème 28.4.5).

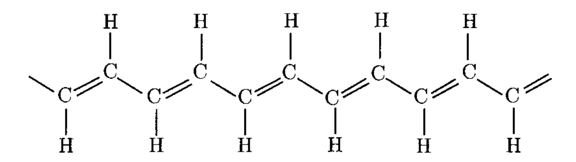


Figure 28.15: Représentation schématique du *trans*-polyacétylène.

On retrouve dans cette longue chaîne les angles de 120°, de sorte que la discussion préliminaire faite à propos du benzène peut être reprise ici, avec les mêmes conclusions : les deux électrons de cœur $(1s)^2$ sont hors course en ce qui concerne le spectre des excitations de nature électronique, ainsi que les trois autres engagés dans les liaisons chimiques, tous ces électrons ne jouant, par leur mouvement rapide, que pour l'existence d'une

[54]Par exemple, avec $N = 12$, on trouve $2|v_1|\cos\frac{\pi}{6} = \sqrt{3}\,|v_1|$.

[55]On a vu dans le Tome I (problème 14.6.11) un exemple où la *densification* du spectre conduit à l'irréversibilité dans la limite d'un nombre infini de fréquences équidistantes.

[56]Pour cette découverte, Shirakawa, Heeger et MacDiarmid ont reçu le Nobel de chimie en 2000.

configuration moléculaire stable : toute déformation coûte de l'énergie mais, d'après la discussion générale de la sous-section 28.1.1, le quantum correspondant est d'un ordre de grandeur plus petit que les excitations électroniques. Ainsi, tout comme pour le benzène, on peut espérer que l'essentiel de la physique de "haute" énergie est bien décrit en retenant comme point de départ un modèle où chaque carbone apporte finalement un seul électron effectif, situé dans l'orbitale $2p_z$; il faudra en temps utile se poser la question du couplage entre ce *gaz* d'électrons et le squelette de la chaîne[57] : on verra qu'il est crucial en raison d'une instabilité sous-jacente qui sera mise en lumière par la suite (voir p. 1297). Clairement, dans ce modèle, le polymère est de fait, topologiquement parlant, un système unidimensionnel ; on note $\mathcal{N}$ le nombre d'atomes de carbone.

Pour les chaînes de longueur modérée, il est avéré expérimentalement depuis longtemps que le squelette est une succession alternée de liaisons longues et courtes – relativement[58] –, ce que l'on représente schématiquement par l'alternance de liaisons simples et doubles. Une question récurrente a été de savoir si cette *dimérisation* de l'ossature de la chaîne persiste dans la limite $\mathcal{N} \gg 1$. Le traitement qui suit répond par l'affirmative [227].

Comme il s'agit d'un très long système, on pourrait être tenté de vouloir profiter de la commodité apportée par des conditions cycliques de Born - von Kármán. On n'en fera rien, laissant la chaîne *ouverte* dans son intégrité topologique. En fin de traitement, on commentera brièvement les modifications – en fait les absences – qui résulteraient d'un tel choix. Par ailleurs, on peut se demander si la *parité* de $\mathcal{N}$ joue un rôle, quoique $\mathcal{N}$ est très grand, puisque le spin total sera entier ou demi-entier selon le cas.

Dans le droit fil du benzène, le traitement minimal est d'introduire un couplage entre premiers voisins, permettant à un électron initialement sur un certain carbone de sauter sur les premiers voisins et pouvoir, de proche en proche, se promener le long de la chaîne tout entière. Compte tenu de l'alternance des liaisons, il faut bien sûr introduire deux éléments de matrice de H différents. Numérotant les atomes de carbone de gauche à droite de 1 à $\mathcal{N}$, et désignant toujours par $|\phi_n\rangle$ l'orbitale $2p_z$ centrée sur le carbone d'indice n, on pose ainsi[59] :

$$\langle\phi_{2n+1}|H|\phi_{2n+2}\rangle = t_2 \ , \qquad \langle\phi_{2n+2}|H|\phi_{2n+3}\rangle = t_1 \quad (n \geq 0) \ , \tag{28.110}$$

t_2 correspondant à une double liaison, t_1 à une simple liaison ; physiquement, on s'attend à $t_2 > t_1$ (la liaison double est plus courte), mais on verra que le cas $t_2 < t_1$ présente un intérêt. Tous les éléments diagonaux de H (ce qui était noté E_0 pour le benzène) sont égaux entre eux, et sont donc omis (cela ne fait que décaler le spectre en bloc). Notant ε la valeur propre, l'équation caractéristique à résoudre traduit l'annulation du

[57] Qu'un tel couplage existe est une évidence physique : l'amplitude de saut d'un électron d'un carbone à l'autre dépend évidemment de la distance entre ces deux atomes. Toute la question est de savoir si cet effet joue un rôle marginal (en tant que petite correction d'énergie, par exemple), ou non.

[58] La différence de longueur n'est que de 0, 08 Å environ.

[59] On adopte ici la notation traditionnelle de la Physique de la matière condensée, où on note traditionnellement t une intégrale de saut entre deux sites.

déterminant $\mathcal{N} \times \mathcal{N}$:

$$\mathcal{D}_{\mathcal{N}}(\varepsilon,\, t_2,\, t_1) \stackrel{\text{déf}}{=} \begin{vmatrix} -\varepsilon & t_2 & 0 & . & . & . & . & . & 0 \\ t_2 & -\varepsilon & t_1 & 0 & . & . & . & . & 0 \\ 0 & t_1 & -\varepsilon & t_2 & 0 & . & . & . & 0 \\ . & 0 & t_2 & -\varepsilon & t_1 & 0 & . & . & 0 \\ . & . & 0 & t_1 & -\varepsilon & t_2 & 0 & . & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & 0 \\ . & . & . & . & . & . & . & -\varepsilon & t_i \\ 0 & 0 & 0 & . & . & . & 0 & t_i & -\varepsilon \end{vmatrix} = 0 \qquad (28.111)$$

Le dernier t_i en bas à droite est t_1 si $\mathcal{N}$ est impair, t_2 si $\mathcal{N}$ est pair. En développant le déterminant suivant la première ligne, on trouve l'égalité :

$$\mathcal{D}_{\mathcal{N}}(\varepsilon,\, t_2,\, t_1) = -\varepsilon\, \mathcal{D}_{\mathcal{N}-1}(\varepsilon,\, t_1,\, t_2) - t_2^2\, \mathcal{D}_{\mathcal{N}-2}(\varepsilon,\, t_2,\, t_1) \ ; \qquad (28.112)$$

afin de simplifier les notations, posons :

$$D_{\mathcal{N}}(\varepsilon) \stackrel{\text{déf}}{=} \mathcal{D}_{\mathcal{N}}(\varepsilon,\, t_2,\, t_1) \ , \quad \tilde{D}_{\mathcal{N}}(\varepsilon) \stackrel{\text{déf}}{=} \mathcal{D}_{\mathcal{N}}(\varepsilon,\, t_1,\, t_2) \ , \qquad (28.113)$$

ce qui permet de récrire (28.112) sous la forme :

$$\begin{cases} D_{\mathcal{N}}(\varepsilon) = -\varepsilon\, \tilde{D}_{\mathcal{N}-1}(\varepsilon) - t_2^2\, D_{\mathcal{N}-2}(\varepsilon) \\ \tilde{D}_{\mathcal{N}-1}(\varepsilon) = -\varepsilon\, D_{\mathcal{N}-2}(\varepsilon) - t_1^2\, \tilde{D}_{\mathcal{N}-3}(\varepsilon) \end{cases} , \qquad (28.114)$$

système pouvant être écrit sous forme matricielle :

$$\begin{bmatrix} 1 & \varepsilon \\ 0 & 1 \end{bmatrix} \begin{bmatrix} D_{\mathcal{N}}(\varepsilon) \\ \tilde{D}_{\mathcal{N}-1}(\varepsilon) \end{bmatrix} = \begin{bmatrix} -t_2^2 & 0 \\ -\varepsilon & -t_1^2 \end{bmatrix} \begin{bmatrix} D_{\mathcal{N}-2}(\varepsilon) \\ \tilde{D}_{\mathcal{N}-3}(\varepsilon) \end{bmatrix} . \qquad (28.115)$$

En introduisant maintenant :

$$\mathbb{D}_{\mathcal{N}} \stackrel{\text{déf}}{=} \begin{bmatrix} D_{\mathcal{N}}(\varepsilon) \\ \tilde{D}_{\mathcal{N}-1}(\varepsilon) \end{bmatrix} , \qquad (28.116)$$

et en multipliant par la matrice inverse de la matrice située à gauche dans (28.115), on obtient :

$$\mathbb{D}_{\mathcal{N}} = \mathbb{M}\mathbb{D}_{\mathcal{N}-2} \ , \qquad \mathbb{M} \stackrel{\text{déf}}{=} \begin{bmatrix} \varepsilon^2 - t_2^2 & \varepsilon t_1^2 \\ -\varepsilon & -t_1^2 \end{bmatrix} . \qquad (28.117)$$

Supposons dans la suite que la chaîne contient $\mathcal{N} = 2N$ monomères ; ce qui précède permet d'écrire formellement la solution comme :

$$\mathbb{D}_{2N} = \mathbb{M}^{N-1}\mathbb{D}_2 \ , \qquad \mathbb{D}_{2N+1} = \mathbb{M}^{N-1}\mathbb{D}_3 \ . \qquad (28.118)$$

On a $D_2 = \begin{vmatrix} -\varepsilon & t_2 \\ t_2 & -\varepsilon \end{vmatrix} = \varepsilon^2 - t_2^2$ et $D_1 = -\varepsilon$, d'où $\mathbb{D}_2 = \begin{bmatrix} \varepsilon^2 - t_2^2 \\ -\varepsilon \end{bmatrix}$. Il suffit maintenant de diagonaliser la matrice $\mathbb{M}$ pour calculer commodément sa puissance N^{e}, et en déduire finalement $\mathbb{D}_{2N}$. L'équation caractéristique de $\mathbb{M}$ est :

$$X^2 + (t_1^2 + t_2^2 - \varepsilon^2)X + (t_1 t_2)^2 = 0 \ , \qquad (28.119)$$

dont les deux racines sont :

$$X_\pm = t_1 t_2\, \mathrm{e}^{\pm \mathrm{i}\theta} \ , \qquad \cos\theta \overset{\text{déf}}{=} \frac{1}{2t_1t_2}(\varepsilon^2 - t_1^2 - t_2^2) \ ; \tag{28.120}$$

noter que l'angle θ n'est pas forcément réel. On introduit maintenant la matrice de passage $\mathbb{P}$ et son inverse, avec $\tau \overset{\text{déf}}{=} \frac{t_1}{t_2}$:

$$\mathbb{P} \overset{\text{déf}}{=} \begin{bmatrix} \varepsilon & \varepsilon \\ 1+\frac{1}{\tau}\mathrm{e}^{-\mathrm{i}\theta} & 1+\frac{1}{\tau}\mathrm{e}^{+\mathrm{i}\theta} \end{bmatrix} \ , \quad \mathbb{P}^{-1} \overset{\text{déf}}{=} \frac{1}{2\mathrm{i}\varepsilon\sin\theta} \begin{bmatrix} \tau + \mathrm{e}^{+\mathrm{i}\theta} & -\varepsilon\tau \\ -(\tau+\mathrm{e}^{-\mathrm{i}\theta}) & \varepsilon\tau \end{bmatrix} \tag{28.121}$$

ce qui permet d'écrire :

$$\mathbb{D} \overset{\text{déf}}{=} \begin{bmatrix} \mathrm{e}^{+\mathrm{i}\theta} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\theta} \end{bmatrix} = \mathbb{P}^{-1}\mathbb{M}\mathbb{P} \quad \Longleftrightarrow \quad \mathbb{M}^N = \mathbb{P}\begin{bmatrix} \mathrm{e}^{+\mathrm{i}N\theta} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}N\theta} \end{bmatrix}\mathbb{P}^{-1} \ . \tag{28.122}$$

Revenant alors à (28.118), on en déduit :

$$\boxed{D_{2N}(\varepsilon) = \frac{(t_1t_2)^N}{\sin\theta}\left[\sin(N+1)\theta + \tau\sin N\theta\right]} \tag{28.123}$$

En définitive, l'équation (28.111) pour un polymère à $\mathcal{N} = 2N$ carbones, $C_{\mathcal{N}}H_{2\mathcal{N}}$, est équivalente à :

$$\boxed{f_N(\theta) \overset{\text{déf}}{=} \frac{\sin(N+1)\theta}{\sin N\theta} = -\tau} \tag{28.124}$$

La fonction $f_N(\theta)$ part de $1+\frac{1}{N}$ en $\theta = 0_+$, diverge pour les θ de la forme entier $\times\, \frac{\pi}{N}$, et vaut par continuité $-1-\frac{1}{N}$ en $\theta = \pi$. En conséquence, si $\tau < 1+\frac{1}{N}$, il y a N intersections réelles pour θ ; si $\tau > 1+\frac{1}{N}$, il y a $N-1$ intersections réelles, et une solution complexe $\tilde{\theta} \overset{\text{déf}}{=} \pi + \mathrm{i}\xi$, où ξ est la solution de l'équation $\sinh(N+1)\xi = \tau\sinh N\xi$. On reviendra par la suite sur la signification physique de cette solution un peu à part – qui constitue un *état de surface.*

Pour une chaîne uniforme (non dimérisée), $t_1 = t_2$, $\tau = 1$; les solutions de (28.124) avec $\tau = 1$ sont $\theta_k(1) = k\frac{2\pi}{2N+1}$. Comme, en toute hypothèse le taux de dimérisation est faible, on peut obtenir une expression approchée de $\theta_k(\tau)$ en développant au voisinage de $\tau = 1$. Dérivant par rapport à τ membre à membre, (28.124) donne :

$$\frac{\mathrm{d}f_N(\theta)}{\mathrm{d}\theta}\frac{\mathrm{d}\theta(\tau)}{\mathrm{d}\tau} = -1 \quad \Longleftrightarrow \quad \left(\frac{\mathrm{d}\theta(\tau)}{\mathrm{d}\tau}\right)_{\tau=1} = -\frac{1}{f_N'\left(\frac{2k\pi}{2N+1}\right)} \tag{28.125}$$

Le calcul donne $f_N'\left(\frac{2k\pi}{2N+1}\right) = -(2N+1)\cot\frac{k\pi}{2N+1}$ d'où :

$$\theta_k(\tau) \simeq \frac{2k\pi}{2N+1} - \frac{(1-\tau)}{2N+1}\tan\frac{k\pi}{2N+1} \ , \quad (k = 1,\, 2,\, ...,\, N) \tag{28.126}$$

Cette expression suppose τ très voisin de 1, et aussi inférieur à $1+\frac{1}{N}$: au-delà du seuil de formation de l'état lié, k varie seulement de 1 à $N-1$, l'état de surface mentionné

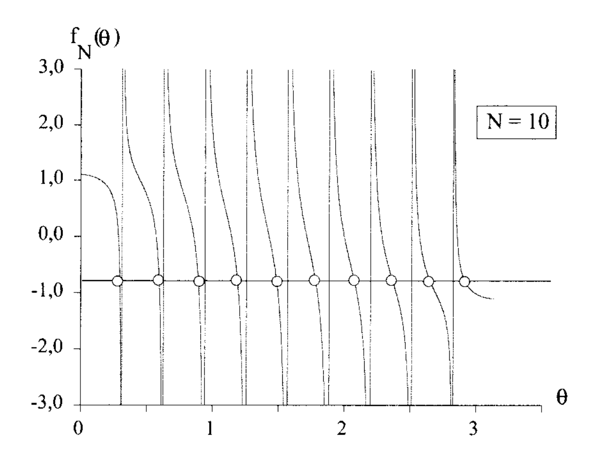


Figure 28.16: Construction graphique des solutions de l'équation (28.124), données par les intersections avec la droite horizontale d'ordonnée $-\tau$. Si τ dépasse $1+\frac{1}{N}$, la solution située à droite devient complexe : c'est l'état de surface décrit dans le texte.

ci-dessus exigeant un traitement spécial. En tout cas, revenant à la définition de l'angle θ apparaissant dans (28.120), on en déduit les valeurs propres de l'énergie dans ce modèle, de part et d'autre de NE_0 qui définit le zéro d'énergie :

$$\boxed{\varepsilon_k(t_1,\, t_2) = \pm\sqrt{t_1^2 + t_2^2 + 2t_1 t_2 \cos\theta_k(\tau)} \qquad (k = 1,\, 2,\, ...,\, N)} \tag{28.127}$$

Pour un polymère à $2N$ atomes de carbone, on obtient ainsi $2N$ niveaux, disposés symétriquement dans les deux intervalles $[-|t_1+t_2|,\, -|t_1-t_2|]$ et $[+|t_1-t_2|,\, +|t_1+t_2|]$, la densité de niveaux augmentant comme N, et forment deux *quasi-bandes* d'énergie quand $N \gg 1$ (voir fig. 28.17). Les deux niveaux les plus proches des deux bandes sont séparés par un *gap* voisin de $2|t_1 - t_2|$ si N est très grand : sa limite à N infini est *finie*.

Pour construire l'état fondamental, il suffit de remplir les N premiers niveaux, chacun avec deux électrons, ce qui revient à saturer la bande inférieure d'énergie. L'état peuplé le plus haut en énergie – plus précisément l'orbitale moléculaire – correspond à $\theta \lesssim \pi$, et a une énergie négative, proche de $-|t_1 - t_2|$. L'énergie de l'état fondamental est simplement donnée par la *somme* des énergies propres, puisque l'on ne tient ici aucun compte de la corrélation électronique[60] :

$$\boxed{E_{\text{fondamental}} = -2\sum_{k=1}^{N} \sqrt{t_1^2 + t_2^2 + 2t_1 t_2 \cos\theta_k(\tau)}} \tag{28.128}$$

[60]... tout comme on l'a déjà fait à propos du benzène. Nul doute que cette omission est une grave carence des modèles ainsi développés.

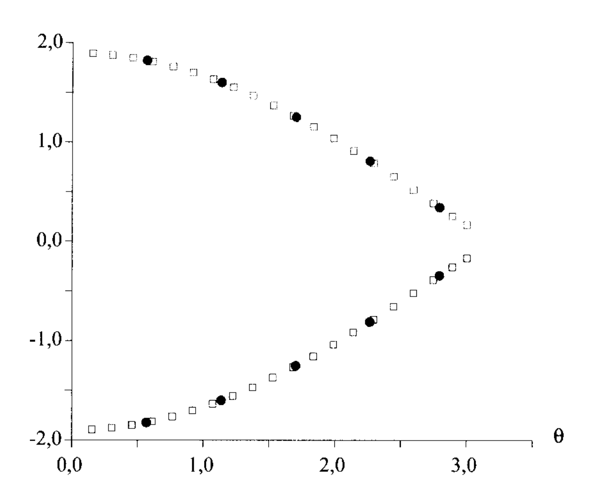


Figure 28.17: Niveaux d'énergie du *trans* - polyacétylène pour $N = 5$ (cercles) et $N = 20$ (carrés) ; le nombre de monomères est égal à $2N$, l'unité d'énergie est t_2. Les courbes sont dessinées avec $\tau = 0,9 < 1 + \frac{1}{N}$ (pas d'état de surface).

Dans le même modèle, la chaîne polymérique *uniforme* (sans alternance des liaisons, $t_1 = t_2 = t$) a pour énergies propres :

$$\varepsilon_k^{(\text{uniforme})} = 2t \cos\frac{k\pi}{2N+1} \ , \quad (k = 1,\, 2,\, ...,\, 2N). \tag{28.129}$$

Dans l'état fondamental, les orbitales sont remplies jusqu'à l'énergie $2|t|\cos\frac{N\pi}{2N+1} \simeq 0$ pour N assez grand, niveau très proche dans les mêmes conditions du premier niveau vide : le *gap* est cette fois très petit, d'autant plus que N est grand (à la limite N infini, il est cette fois *nul*). L'énergie vaut :

$$\boxed{E_{\text{fondamental}}^{(\text{uniforme})} = -2|t|\,\frac{\sin\frac{N\pi}{2(2N+1)}}{\sin\frac{\pi}{2(2N+1)}}\,\cos\frac{(N+1)\pi}{2(2N+1)}} \tag{28.130}$$

Dans la limite $N \gg 1$, on a :

$$E_{\text{fondamental}}^{(\text{uniforme})} = -N\frac{4}{\pi}|t| + \big(1 - \frac{2}{\pi}\big) + \mathcal{O}(N^{-1}) \tag{28.131}$$

Il est intéressant de noter que l'énergie est visiblement *extensive* dans cette limite, une propriété prévisible par le fait que les interactions entre atomes de carbone sont à *courte* portée (la même extensivité apparaît bien sûr aussi pour l'expression (28.128) : la substitution $\sum_k \rightarrow \frac{L}{2\pi}\int \mathrm{d}k$ dans la limite $N \gg 1$ fait ressortir en facteur la longueur $L \equiv 2Na$ de la chaîne, et introduit la fonction elliptique complète de deuxième espèce $\mathbf{E}$, définie en (28.222)).

Ces considérations laissent pressentir que l'énergie électronique du *gaz* d'électrons est plus basse quand la chaîne est alternée, chaque carbone étant à l'abscisse $na+(-1)^n u$,

$0 \leq u \ll a$, ce qui semble indiquer une *instabilité* du système[61] dès que l'on autorise le réseau à se déformer. Le calcul détaillé confirme cette intuition, et montre que la chaîne unidimensionnelle habillée avec ses électrons a en effet tendance à se dimériser spontanément ; hors toute autre compétition énergétique, le système serait bel et bien instable.

En réalité, le bilan énergétique doit inclure le coût d'énergie de déformation du squelette moléculaire. Celui-ci est réputé avoir une configuration d'équilibre, correspondant au minimum d'une certaine énergie potentielle, de sorte que chaque liaison ne peut être allongée ou racourcie sans dépenser de l'énergie. Pour savoir qui l'emporte de ces deux tendances antagonistes, il faut, par exemple examiner la courbure (la concavité) de l'énergie *totale* en $u = 0$, et voir si elle est positive ou négative. En ce qui concerne la déformation du squelette, il est évident qu'il y a un coût par liaison, de sorte que le coût total est proportionnel au nombre de monomères dans la chaîne. Quant au gain en énergie venant du gaz d'électrons, il est plus ardu à déterminer, mais on peut montrer qu'il donne une concavité (négative évidemment) variant en gros comme $N \ln N$ et qui, de ce fait, l'emporte toujours sur la dépense en énergie élastique. En conséquence, la concavité est *négative* en $u = 0$: la chaîne se déforme jusqu'à ce qu'un nouveau point d'équilibre se trouve pour une certaine valeur u_0, donnant une chaîne dimérisée. En fonction de l'écart à la chaîne uniforme, l'énergie totale (gaz d'électrons + chaîne élastique) a donc l'allure d'un double puits (u et $-u$ donnent la même énergie totale) : cette image résumant la situation renvoie également à la double dégénérescence des deux structures, chacune miroir de l'autre, mentionnée tout au début. Tous ces détails sont examinés dans le problème 28.4.5, où on discute également le rôle de la *corrélation* entre les électrons, nullement prise en compte dans le traitement précédent.

Pour terminer, disons un mot de l'état de surface, qui apparaît en tant que solution complexe pour θ de l'équation (28.124). Sans qu'il soit utile de les écrire en détail, on se doute que les orbitales moléculaires $|\psi_k\rangle$ associées aux énergies propres ε_k contiendront des facteurs de phase du genre $\mathrm{e}^{\mathrm{i}kn\theta}$, donnant des densités électroniques diffuses et délocalisées sur toute la chaîne (comme pour le benzène) quand l'angle θ est réel (c'est pourquoi on parle d'états *étendus*). Au contraire, quand θ est complexe – et une telle solution existe quand $\tau > 1 + \frac{1}{N}$ –, les mêmes facteurs deviennent des exponentielles décroissantes quand on s'éloigne de l'extrémité de la chaîne, et c'est pour cette raison que l'on parle alors d'état *de surface*. Un tel état n'aurait bien évidemment pas pu apparaître si l'on avait utilisé des conditions cycliques, et c'est pourquoi, en ce sens et même si la longueur de la chaîne est arbitrairement grande, il y a une différence qualitative entre les deux traitements, avec les vraies conditions aux limites (franches) ou avec des conditions cycliques (chaîne refermée sur elle-même). Physiquement, cette différence est d'autant plus importante que cette valeur complexe de θ donne une énergie (réelle !) qui se trouve *dans le gap* séparant les deux bandes d'énergie. En effet, posant $\tilde{\theta} = \pi + \mathrm{i}\xi$ (donnant $\cos\tilde{\theta} = -\cosh\xi$), l'équation à résoudre est $\sinh(N+1)\xi = \tau \sinh N\xi$, qui donne un certain ξ_0. Selon la définition de θ, (28.120), l'énergie correspondante est :

$$\varepsilon_{\text{surface}}(t_1,\, t_2) = \pm\sqrt{t_1^2 + t_2^2 - 2t_1 t_2 \cosh \xi_0(\tau)} \ . \qquad (28.132)$$

[61] C'est l'instabilité de Peierls, que l'on retrouvera dans le chapitre 30.

Quelle que soit la valeur précise du nombre ξ_0, ces deux énergies sont en module plus petites que $|t_1 - t_2|$: elles sont bien dans le *gap* d'énergie entre les deux bandes. L'apparition d'états localisés dans un *gap* est un phénomène courant, et très important physiquement.

28.2.6 Nature physique de la liaison chimique

La nature de la liaison chimique résulte clairement d'une compétition entre deux interactions antagonistes : les noyaux, ou les cœurs, se repoussent, mais chacun d'entre eux attire les mêmes électrons ; ceux-ci, par leur va-et-vient, jouent le rôle d'une *colle* entre des particules qui, nues, ont tendance à se rejeter mutuellement. Cette image est universelle : toute interaction entre deux systèmes, fondamentale ou comme ici effective, résulte de l'échange de particules entre ces deux systèmes.

Les cœurs coopèrent en attirant chacun les électrons et ce sont ces efforts conjugués qui assurent, éventuellement, la stabilité de l'édifice. Il s'agit bien d'ingrédients *nécessaires* ; l'existence ou non d'une liaison stable résulte évidemment du fait que l'interaction effective induite par les allers-retours des électrons est assez forte pour prendre le dessus et, tout compte fait, donner à l'énergie $E(Q_j)$ un *minimum.* C'est la nature même de la distribution électronique qui conditionne l'existence d'une liaison, laquelle trouve son origine, comme on va le voir, dans la compétition subtile entre les différentes contributions à l'énergie. Pour la simplicité des choses, on raisonne dans la suite avec une molécule diatomique, dont les noyaux sont à la distance R l'un de l'autre.

Le théorème du Viriel permet d'analyser finement cette compétition, en fournissant des relations simples entre les valeurs moyennes des énergies cinétique et potentielle ; sa démonstration consiste à manipuler la quantité $\vec{r}.\vec{p}$ (pour un seul électron), tout comme en Mécanique classique on s'intéresse, pour la démonstration analogue, à la moyenne temporelle de $\vec{r}.\vec{p}$ [51] ; le commutateur de cette quantité avec $H_0 = T + V$ est :

$$[H_0, \vec{r}.\vec{p}] = \sum_{u=x,\,y,\,z} [H_0, up_u] = \sum_{u=x,\,y,\,z} ([H_0, u]\,p_u + u[H_0, p_u]) \quad . \tag{28.133}$$

Compte tenu de :

$$[H_0, u] = \frac{1}{2m}\,[\vec{p}^{\,2}, u] = -\mathrm{i}\frac{\hbar}{m}p_u \ , \qquad [H_0, p_u] = [V, p_u] = +\mathrm{i}\hbar\frac{\partial V}{\partial u} \tag{28.134}$$

il vient :

$$[H_0, \vec{r}.\vec{p}] = \sum_{u=x,\,y,\,z} \left(-\mathrm{i}\frac{\hbar}{m}p_u^2 + \mathrm{i}\hbar u\frac{\partial V}{\partial u}\right) = \mathrm{i}\hbar(-2T + \vec{r}.\vec{\nabla}V) \ . \tag{28.135}$$

Par ailleurs, si toutes les interactions sont de Coulomb, V est une fonction homogène de degré -1 :

$$V(\vec{r}, \vec{R}\,) = f(x, y, z, R) \quad \text{avec : } f(\lambda x, \lambda y, \lambda z, \lambda R) = \frac{1}{\lambda}\,f(x, y, z, R) \ ; \tag{28.136}$$

la dérivation en λ de la dernière égalité donne :

$$x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}+z\frac{\partial f}{\partial z}+R\frac{\partial f}{\partial R}=-\frac{1}{\lambda^2}f \tag{28.137}$$

et si on fait $\lambda = 1$ partout, il vient $x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}+z\frac{\partial f}{\partial z}+R\frac{\partial f}{\partial R}=-f(x,\,y,\,z,\,R)$: c'est la formule d'Euler pour une fonction homogène de degré -1. Pour $V(\vec{r},\,\vec{R})$ le premier membre de (28.137) est $\vec{r}.\vec{\nabla}f+R\frac{\partial f}{\partial R}$, d'où l'égalité :

$$\vec{r}.\vec{\nabla}V+R\frac{\partial V}{\partial R}=-V\ , \tag{28.138}$$

qui permet d'écrire le commutateur (28.135) sous la forme :

$$\boxed{[H_0,\,\vec{r}.\vec{p}]=\mathrm{i}\hbar\Big(-2T-V-R\frac{\partial V}{\partial R}\Big)} \tag{28.139}$$

Tout ceci reste vrai pour un nombre N quelconque d'électrons : il suffit de manipuler la somme $\sum_{i=1}^{N}\vec{r}_i.\vec{p}_i$. Maintenant :

1. la valeur moyenne d'un commutateur $[H_0,\,\Omega]$ est nulle dans tout état propre normalisable $|E\rangle$ de H_0 :

$$\langle E|[H_0,\,\Omega]|E\rangle=(E-E)\,\langle E|\Omega|E\rangle=0\ , \tag{28.140}$$

la dernière égalité supposant que $\langle E|\Omega|E\rangle$ est une quantité finie, ce qui est vrai pour tout état lié et pour toute observable suffisamment régulière[62]. Il résulte alors de (28.139) que la moyenne prise sur tout état électronique normalisable satisfait :

$$\langle 2T+V+R\frac{\partial V}{\partial R}\rangle=0\ . \tag{28.141}$$

2. Par ailleurs, soit un Hamiltonien dépendant d'un paramètre ζ, $H(\zeta)$; ses valeurs propres E et ses fonctions propres ψ_E, dépendent aussi de ce paramètre, soit $E(\zeta)$ et $\psi_E(\zeta)$. Comme, pour une fonction normalisable, $\frac{\mathrm{d}}{\mathrm{d}\zeta}\langle\psi_E(\zeta)|\psi_E(\zeta)\rangle=0$, il vient :

$$\frac{\mathrm{d}E}{\mathrm{d}\zeta}\equiv\frac{\mathrm{d}}{\mathrm{d}\zeta}\langle\psi_E(\zeta)|H(\zeta)|\psi_E(\zeta)\rangle=\langle\psi_E(\zeta)|\frac{\partial H}{\partial\zeta}|\psi_E(\zeta)\rangle\equiv\langle\frac{\partial H}{\partial\zeta}\rangle\ , \tag{28.142}$$

d'où l'égalité exprimant le théorème de Hellmann - Feynman :

$$\boxed{\frac{\mathrm{d}E}{\mathrm{d}\zeta}=\langle\frac{\partial H}{\partial\zeta}\rangle} \tag{28.143}$$

[62]On entend par là que la moyenne de Ω existe ; ceci n'est pas assuré pour des observables mêmes relativement *banales* : à propos de la structure hyperfine (voir chapitre 27, section 27.4), on a rencontré le cas de r^{-3}, dont la moyenne n'existe pas dans les états S (l'intégrale diverge logarithmiquement à l'origine). On peut aussi dire : quand H_0 et Ω sont hermitiques, leur commutateur est anti-hermitique ; sa valeur moyenne est donc de la forme i×nombre réel, un nombre qui ne peut être que zéro, pourvu qu'il soit fini.

Noter que toutes les moyennes résultent d'une intégration exclusive sur $\vec{r}$ et restent donc des fonctions de R.

Avec $H \equiv H_0$ et $\zeta \equiv R$, (28.143) donne :

$$\frac{\mathrm{d}E}{\mathrm{d}R} = \langle \frac{\partial H_0}{\partial R} \rangle = \langle \frac{\partial V}{\partial R} \rangle \ , \tag{28.144}$$

ce qui permet de récrire (28.141) sous la forme :

$$\langle 2T \rangle + \langle V \rangle + R\frac{\mathrm{d}E}{\mathrm{d}R} = 0 \ , \tag{28.145}$$

où V désigne toujours l'énergie potentielle totale. Par ailleurs, on a évidemment :

$$E = \langle T \rangle + \langle V \rangle \ , \tag{28.146}$$

d'où, par élimination, soit de $\langle V \rangle$, soit de $\langle T \rangle$ entre (28.145) et (28.146), les deux égalités :

$$\boxed{\langle T \rangle = -E - R\frac{\mathrm{d}E}{\mathrm{d}R} \qquad \langle V \rangle = 2E + R\frac{\mathrm{d}E}{\mathrm{d}R}} \tag{28.147}$$

Ces expressions permettent de montrer que l'existence d'une liaison chimique stable résulte de la *diminution* de l'énergie potentielle totale et d'une *augmentation* de l'énergie cinétique relativement à leurs valeurs pour la molécule dissociée ($R = \infty$). En effet, s'il existe une position d'équilibre R_0, alors, en ce point (où la dérivée $\frac{\mathrm{d}E}{\mathrm{d}R}$ s'annule) :

$$\langle T \rangle_{R_0} = -E(R_0) \qquad \langle V \rangle_{R_0} = 2E(R_0) \ . \tag{28.148}$$

Par ailleurs, pour $R = +\infty$, on a aussi[63] :

$$\langle T \rangle_{\infty} = -E(\infty) \qquad \langle V \rangle_{\infty} = 2E(\infty) \ . \tag{28.149}$$

D'où :

$$\boxed{\langle T \rangle_{R_0} - \langle T \rangle_{\infty} = E(\infty) - E(R_0) > 0 \qquad \langle V \rangle_{R_0} - \langle V \rangle_{\infty} = 2[E(R_0) - E(\infty)] < 0} \tag{28.150}$$

Physiquement, l'augmentation d'énergie cinétique lors de la formation de la liaison se comprend bien : initialement localisés près d'un cœur pour les très grands R, les électrons engagés dans une liaison peuvent en fait aller sur l'autre quand R devient assez petit. Ces excursions se traduisent naturellement par une augmentation de l'énergie cinétique, comme on l'a d'ailleurs vu à propos du benzène. Toutes choses égales par ailleurs, la liaison est donc d'autant plus forte que les électrons vont *vite* d'un cœur à l'autre.

Enfin, en terme de densité électronique, les exemples traités montrent que, tout naturellement, la liaison stable (état liant) contient une accumulation d'électrons entre les cœurs, alors qu'une liaison instable (état antiliant) présente un déficit d'électrons entre ceux-ci, un fait dont la figure 28.18 est une illustration.

[63] On suppose que $\lim_{R\to+\infty} R\,\frac{\mathrm{d}E}{\mathrm{d}R} = 0$, ce qui suppose $\alpha > 0$ si $E(R) \sim E_\infty + CR^{-\alpha}$ (on exclut par exemple une dépendance logarithmique $\sim [\ln \frac{R}{a_0}]^{-1}$).

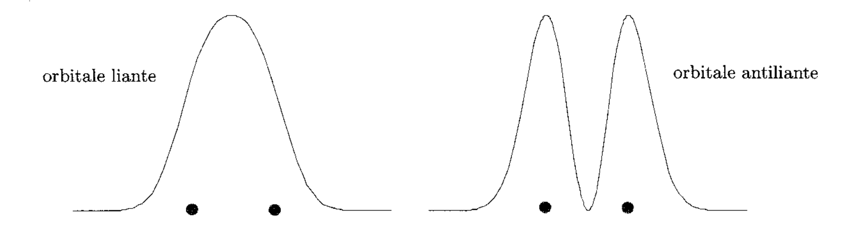


Figure 28.18: Illustration schématique de la différence de densité électronique associée à deux orbitales l'une liante, l'autre antiliante.

28.3 Mouvement des noyaux. Spectres de vibration et de rotation

L'analyse de la structure électronique permet de comprendre l'origine de la liaison chimique et, plus généralement, fournit un schéma explicatif de la stabilité des molécules. De surcroît, la mise en évidence des états électroniques – où les noyaux sont fixes par définition – permet de comprendre l'existence des spectres UV et visible des molécules – à condition toutefois de ne pas rentrer trop dans le détail[64].

La description précise du mouvement des noyaux (une fois moyenné celui des électrons) est le fondement théorique d'une autre spectroscopie (IR et micro-ondes[65]) apportant de précieux renseignements sur la conformation moléculaire et les champs de forces effectifs entre noyaux. Principalement, la spectroscopie IR (spectres de vibration) donne des renseignements sur les constantes de force d'équilibre et donc sur la *force* des liaisons chimiques, tandis que la spectroscopie micro-ondes permet de voir les états rotationnels, la molécule étant considérée (en première approximation) comme un rotateur rigide (pas de vibrations), défini par des axes principaux et des moments d'inertie ; connaissant les masses, la détermination de ces derniers fournit notamment les distances d'équilibre des liaisons. Au total, les mesures spectroscopiques mettant en jeu les degrés de libertés vibrationnels et rotationnels permettent de remonter à la configuration géométrique de la molécule et aux forces effectives qui la régissent.

Pour une molécule à N noyaux, il y a $3N$ degrés de liberté nucléaires. Comme on l'a vu, seuls $3N - 6$ d'entre eux – ou $3N - 5$ pour une molécule linéaire[66] – correspondent à des mouvements internes (c'est-à-dire produisant des *déformations* du squelette moléculaire). Parmi les autres degrés, ceux de translation ne donnent jamais de quantification (translation en bloc dans $\mathbb{R}^3$ sans conditions aux limites) ; en revanche, la rotation en bloc est forcément quantifiée, en conséquence du fait que les variables dynamiques sont

[64] Par exemple, toute raie d'émission s'accompagne de satellites provenant de transitions électroniques avec modification de l'état de vibration intramoléculaire.

[65] La spectroscopie Raman, très utile pour les molécules dénuées de moment dipolaire électrique – voir ci-dessous – met en jeu des phénomènes généralisant la diffusion Rayleigh classique, et utilise donc une source d'excitation de fréquence optique, typiquement.

[66] On entend par là une molécule dont la configuration d'équilibre est linéaire et qui ne peut se déformer transversalement, cas-limite où les constantes de torsion associées sont très grandes.

des *angles*, définis sur un intervalle *fini*, $[0,\, 2\pi]$ ou $[0,\, \pi]$ suivant les cas. En Mécanique quantique, toute variable confinée produit la quantification *via* le jeu des conditions aux limites imposées par le sens physique accordé aux fonctions d'onde. En conséquence de la théorie du moment cinétique, la fonction est inchangée par une rotation de 2π des variables *spatiales* puisqu'il n'est ici question que d'un moment cinétique *orbital.*

L'étude détaillée des mouvements nucléaires est assez laborieuse pour une molécule un tant soi peu complexe[67] ; on se bornera principalement ici à exposer les idées essentielles, données ci-dessous dans le cas le plus simple, celui des molécules diatomiques.

28.3.1 Molécules diatomiques

Dans cette sous-section, on considère exclusivement le cas d'une molécule AB, hétéronucléaire ($\mathrm{A} \neq \mathrm{B}$) ou homonucléaire (A=B). Dans cette seconde étape de l'approximation adiabatique, on dispose d'une énergie électronique E_m, jouant le rôle d'une énergie potentielle effective dans une équation aux valeurs propres pour (ici) deux noyaux ponctuels de masses M_A et M_B (voir (28.33)). E_m ne dépend ici que d'une seule coordonnée, la distance R entre les deux noyaux (homogénéité et isotropie galiléennes de l'espace). S'agissant d'étudier les mouvements de la molécule constituée, la fonction $E_m(R)$ est supposée posséder un minimum (unique, pour simplifier) en $R = R_0$.

Fonctions propres de vibration - rotation

L'équation nucléaire a la forme explicite :

$$\Big[-\frac{\hbar^2}{2M_\mathrm{A}}\,\Delta_\mathrm{A} - \frac{\hbar^2}{2M_\mathrm{B}}\,\Delta_\mathrm{B} + E_m(R)\Big]\,\chi_m(\vec{R}_\mathrm{A}, \vec{R}_\mathrm{B}) = E\,\chi_m(\vec{R}_\mathrm{A}, \vec{R}_\mathrm{B}) \ . \tag{28.151}$$

Ceci représente l'équation aux valeurs propres standard pour deux particules en interaction centrale par $E_m(R)$. La séparation du centre de masse fait apparaître χ_m sous la forme :

$$\chi_m(\vec{R}_\mathrm{A}, \vec{R}_\mathrm{B}) = \mathrm{e}^{\mathrm{i}\,\vec{K}.\vec{X}}\,F_m(\vec{R}) \qquad (\vec{R} = \vec{R}_\mathrm{B} - \vec{R}_\mathrm{A}) \ , \tag{28.152}$$

où $\vec{X}$ est la position du centre de masse :

$$\vec{X} = \frac{M_\mathrm{A}\vec{R}_\mathrm{A} + M_\mathrm{B}\vec{R}_\mathrm{B}}{M_\mathrm{A} + M_\mathrm{B}} \ , \tag{28.153}$$

et le vecteur d'onde $\vec{K}$ est relié à l'énergie de translation en bloc par :

$$E_\mathrm{translation} = \frac{\hbar^2\vec{K}^{\,2}}{2(M_\mathrm{A} + M_\mathrm{B})} \ . \tag{28.154}$$

[67]Quelques indications sont données dans la section 28.3.3.

Par ailleurs, la fonction F_m satisfait ($\mu^{-1} = M_{\rm A}^{-1} + M_{\rm B}^{-1}$) :

$$\Big[-\frac{\hbar^2}{2\mu}\Delta_{\vec R} + E_m(R)\Big]F_m(\vec R) = \Big(E - \frac{\hbar^2\vec K^{\,2}}{2M}\Big)F_m(\vec R) \ . \tag{28.155}$$

μ étant une masse (réduite) *nucléaire*, le terme cinétique est petit par rapport à la profondeur du puits de potentiel que constitue $E_m(R)$. Dans le repère du centre de masse ($\|\vec K\| = 0$) – le seul pertinent ici –, on a :

$$\Big[-\frac{\hbar^2}{2\mu}\Delta_{\vec R} + E_m(R)\Big]F_m(\vec R) = EF_m(\vec R) \ . \tag{28.156}$$

Ceci fait renouer avec un cadre familier : (28.156) est une équation aux valeurs propres pour une particule de masse μ dans le potentiel radial $E_m(R)$. La séparation des variables angulaires procède donc exactement comme pour l'atome d'hydrogène et fait apparaître les harmoniques sphériques, que l'on notera ici Y_{LM}. En désignant par $(R,\,\theta,\,\phi)$ les variables sphériques conventionnelles (voir fig. 28.19), on pose[68] :

$$F_m(\vec R) = \frac{1}{R}\,f(R)\,Y_{LM}(\theta,\phi) \qquad (L \in \mathbb{N}\ ,\ M = -L,\,-L+1,\,\ldots,\,+L)\ , \tag{28.157}$$

et la fonction radiale f satisfait :

$$-\frac{\hbar^2}{2\mu}\,\frac{\mathrm{d}^2 f}{\mathrm{d}R^2} + V_{\rm eff}(R)f(R) = Ef(R) \qquad (R > 0)\ , \tag{28.158}$$

avec :

$$V_{\rm eff}(R) = E_m(R) + \frac{\hbar^2 L(L+1)}{2\mu R^2} \ . \tag{28.159}$$

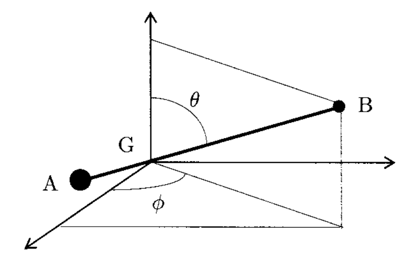


Figure 28.19: Repère utilisé pour fixer la position d'une molécule diatomique AB.

Comme d'habitude pour un champ central (voir (19.23)), le potentiel effectif est la somme de l'énergie potentielle purement radiale et du terme centrifuge $\frac{\vec L^{\,2}}{2\mu R^2}$, provenant de l'opérateur cinétique $T_{\rm N}$. De toute évidence, la variable R décrit la vibration de la

[68] L'indice m de F_m continue à désigner un ensemble de nombres quantiques. Par ailleurs, la fonction radiale f dépend explicitement de L, mais ceci est omis par simplicité.

molécule, alors que les variables angulaires sont associées à sa rotation. Il est utile à ce stade de mettre des nombres dans (28.159), ce qui permet de réaliser que les deux contributions à V_{eff} ne sont pas du même ordre de grandeur. En effet (m est la masse de l'électron, μ une masse réduite nucléaire) :

$$\frac{\hbar^2 L(L+1)}{2\mu R^2} \sim \frac{\hbar^2}{2\mu(2a_0)^2} = \frac{\hbar^2}{8\mu(\frac{\hbar^2}{me'^2})^2} = \frac{m}{\mu}\,\frac{me'^4}{8\hbar^2} = \frac{m}{4\mu}\,E_{\text{I}} \ , \tag{28.160}$$

d'une part. D'autre part, $E_m \sim E_{\text{I}}$, de sorte que :

$$\frac{\hbar^2 L(L+1)}{2\mu R^2} \sim \frac{m}{4\mu}\,E_m \ll E_m \ ; \tag{28.161}$$

sans surprise, l'énergie de rotation est bien une petite correction devant l'énergie E_m ; en première approximation, $V_{\text{eff}}(R) \simeq E_m(R)$, quel que soit R.

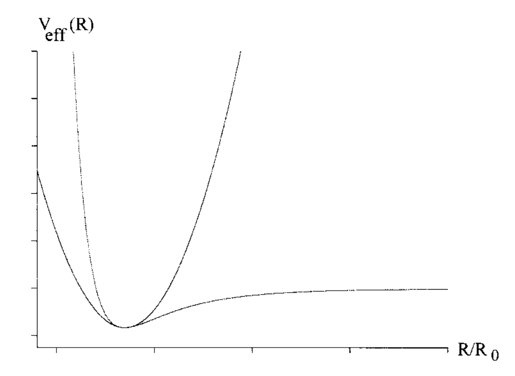


Figure 28.20: Approximation de $V_{\text{eff}}(R)$ par sa parabole osculatrice (même minimum, même courbure au minimum).

Par ailleurs, ce sont les vibrations de la molécule constituée qui présentent un intérêt, en tant que source d'information sur l'édifice moléculaire tel qu'il existe. Implicitement, ceci suppose que les vibrations sont de petite amplitude (autrement, la liaison chimique se casse). Il est donc licite, pour la question posée – mais à condition, dans la suite, de ne pas impliquer des niveaux vibrationnels trop excités –, de développer $V_{\text{eff}}(R)$ près de son minimum et d'écrire[69] :

$$V_{\text{eff}}(R) \simeq E_m(R_0) + \frac{1}{2}\,E_m''(R_0)\,(R-R_0)^2 + \frac{\hbar^2 L(L+1)}{2\mu R_0^2} - \frac{\hbar^2 L(L+1)}{\mu R_0^3}\,(R-R_0) \ . \tag{28.162}$$

Le troisième terme dans (28.162) est l'énergie du rotateur rigide, la distance R étant fixée à sa valeur d'équilibre R_0. Le dernier terme, qui est une petite correction d'après

[69] R_0 dépend de l'état électronique considéré ; on devrait donc plutôt noter R_{m0} la valeur d'équilibre, ce que l'on ne fait pas pour alléger l'écriture.

les estimations précédentes, est visiblement un couplage entre vibration et rotation : sous l'effet de celle-ci, la liaison s'étire un peu et le point d'équilibre change (un peu), sans pour autant changer la fréquence[70] de fond de puits (qui ne dépend que de la dérivée seconde). Dans la suite, oubliant cette correction et se bornant au terme harmonique provenant de $E_m(R)$, l'équation à résoudre, (28.158), prend la forme simplifiée :

$$-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2 f}{\mathrm{d}R^2} + \Big[E_m(R_0) + \frac{1}{2}\mu\omega_m^2(R-R_0)^2 + B_m hL(L+1)\Big] f(R) = E\,f(R) \ , \quad (28.163)$$

où ont été introduites la pulsation de vibration ω_m et la notation traditionnelle pour le terme centrifuge[71] :

$$\boxed{\mu\omega_m^2 = E''_m(R_0) \qquad B_m = \frac{\hbar}{4\pi\mu R_0^2} \equiv \frac{\hbar}{4\pi I_0}} \quad (28.164)$$

I_0 désignant le moment d'inertie pour la distance d'équilibre R_0.

Clairement, on peut attendre des effets isotopiques importants, exclusivement dus au terme cinétique T_{N} : par exemple, entre l'hydrogène H_2 et son isotope D_2, on gagne un facteur 2 pour la masse réduite μ. D'un autre côté, la dérivée seconde $E''_m(R_0)$ ne dépend pas des masses nucléaires[72] : la variation de masse réduite se reporte donc intégralement sur les pulsations. En raison de la première relation (28.164), la pulsation doit être divisée par $\sqrt{2}$ quand on passe de H_2 à D_2, et c'est bien ce que l'on observe[73] : la vibration strictement harmonique se produit[74] à 4401 cm^{-1} pour H_2, et à 3112 cm^{-1} pour D_2. D'une façon générale, la pulsation ω_m est d'autant plus grande que la masse réduite est faible et la liaison stable (le minimum de $E_m(R)$ est aigu). Pour une molécule de gaz rare (en interaction par van der Waals), $\hbar\omega_m$ est beaucoup plus petit (quelques dizaines de cm^{-1} au plus), traduisant la faiblesse (et la fragilité) d'une liaison de type moléculaire.

L'équation (28.163) est une approximation de (28.158), et représente un oscillateur harmonique linéaire[75] pour l'écart $R - R_0$, dont le spectre est décalé de la constante $B_m h\,L(L+1)$. D'après ce que l'on sait de l'oscillateur harmonique linéaire, on peut dès lors affirmer que la valeur propre E de (28.163) est de la forme :

$$\boxed{E = E_m(R_0) + \hbar\omega_m(v + \frac{1}{2}) + B_m h\,L(L+1) \qquad (v \in \mathbb{N},\ L \in \mathbb{N})} \quad (28.165)$$

Cette expression met bien en évidence le découplage des différents degrés de liberté, l'énergie apparaissant sous la forme d'une *somme* de trois termes, écrits dans l'ordre des énergies décroissantes :

$$E = E_{\mathrm{el}} + E_{\mathrm{vibr}} + E_{\mathrm{rot}} \ . \quad (28.166)$$

[70]Le terme suivant du développement de l'énergie centrifuge est en $(R - R_0)^2$ et "renormalise" donc la fréquence de vibration.

[71]La constante B_m dépend de l'état électronique considéré *via* la distance d'équilibre R_0 de la liaison ; B_m est homogène à une fréquence.

[72]L'énergie $E_m(R)$ se trouve dans l'étape où les noyaux sont fixes, et est donc forcément la même pour tous les isotopes.

[73]$4401/3112 \simeq 0,999993 \times \sqrt{2}$.

[74]1 eV$=\frac{|e|}{100hc}$ cm$^{-1} \simeq 8\,065,55\,\mathrm{cm}^{-1}$; $1\,\mathrm{cm}^{-1} \simeq 30\,\mathrm{GHz}$.

[75]À une petite réserve près liée au fait qu'ici R est une variable positive – voir plus loin.

Compte tenu des estimations des énergies de vibration et de rotation, $B_m \ll \frac{\omega_m}{2\pi}$, de sorte que les niveaux de vibration et de rotation sont disposés comme illustré sur la figure 28.21.

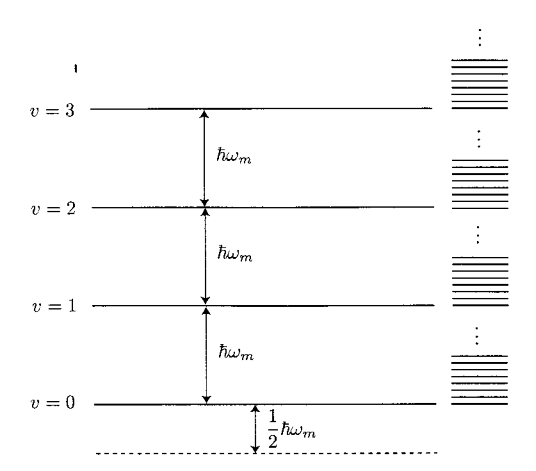


Figure 28.21: Représentation schématique des niveaux de vibration et de rotation d'une molécule diatomique. Le trait en pointillés situe le zéro d'énergie. Les niveaux de vibration ont l'énergie $\hbar\omega_m(v + \frac{1}{2})$, $v \in \mathbb{N}$; les peignes très fins schématisent les niveaux de rotation, décalés de $hB_m L(L+1)$ par rapport aux niveaux de vibration pure.

En ce qui concerne la fonction propre f, c'est – formellement – une fonction propre d'oscillateur harmonique centrée en R_0, à ceci près que la variable R est ici cantonnée aux valeurs positives – donc la variable $R - R_0$ donnant l'écart à la position d'équilibre varie seulement de $-R_0$ à $+\infty$, et non entre $\pm\infty$. Ce *distinguo* n'est pas gênant en pratique, puisque toutes les fonctions d'onde décroissent très vite de part et d'autre de R_0 (l'enveloppe est gaussienne). Les différences deviennent sensibles seulement pour les grandes valeurs du nombre quantique de vibration v, mais, de toute façon, l'hypothèse des petites oscillations exclut la considération d'états de vibration très excités, tout comme l'approximation harmonique d'ailleurs. Plus précisément, les fonctions propres harmoniques sont localisées autour de R_0 sur un intervalle ΔR ; pour que la queue gaussienne du côté $R < 0$ soit invisible et donc inoffensive, il faut $\Delta R \ll R_0$. Comme $\Delta R \sim (\frac{\hbar}{\mu\omega_m})^{1/2}$, il faut avoir :

$$\sqrt{\frac{\hbar}{\mu\omega_m}} \ll R_0 \Longleftrightarrow \sqrt{\frac{\hbar}{\mu\sqrt{E''_m(R_0)/\mu}}} \ll R_0, \Longleftrightarrow \frac{\hbar^2}{\mu E''_m(R_0)R_0^2} \ll R_0^2 \ . \qquad (28.167)$$

Par ailleurs, $E''_m(R_0)R_0^2 \sim \frac{\hbar^2}{ma_0^2}$; en reportant cette estimation dans (28.167), on trouve que (28.167) est équivalente à $\frac{m}{\mu} \ll 1$, une inégalité largement satisfaite puisque μ est une masse réduite d'origine *nucléaire*.

Au total, et quand on néglige le couplage vibration - rotation, les fonctions propres

ont la forme factorisée :

$$\boxed{F_{vLM} = \frac{1}{R}\, f_v(R)\, Y_{LM}(\theta, \phi) \equiv u_v(R)\, Y_{LM}(\theta, \phi)} \tag{28.168}$$

$f_v(R)$ est une gaussienne multipliée par un polynôme de Hermite de degré v, H_v :

$$f_v(R) = C_v\, H_v\Big(\sqrt{\frac{\mu\omega_m}{\hbar}}\,(R - R_0)\Big)\, \mathrm{e}^{-\frac{\mu\omega_m}{2\hbar}(R-R_0)^2} \quad , \tag{28.169}$$

où C_v est la constante de normalisation (voir son expression, Tome I, éq. (16.74)). Y_{LM} est une harmonique sphérique qui décrit la rotation du bâton rigide que constitue la molécule diatomique avec ses noyaux (cœurs) fixés à la distance R_0, dont les seuls degrés de liberté sont les deux angles θ et ϕ.

Spectres de vibration et de rotation

À l'ordre le plus bas, l'interaction entre un système microscopique et le champ électromagnétique classique est du genre $-\vec{d}.\vec{\mathcal{E}}$, où $\vec{\mathcal{E}}$ est le champ électrique de l'onde[76] et où $\vec{d}$ est le moment dipolaire électrique total[77] de la molécule (neutre) moyenné sur le mouvement électronique associé à l'état électronique considéré ; s'agissant ici d'examiner les transitions entre états dynamiques des noyaux, on raisonne à état électronique donné, ne considérant que les degrés de liberté nucléaires. $\vec{d}$ est donc :

$$\vec{d} = \langle \Phi_m | \sum_i e\vec{r}_i + \sum_j Z_j |e| \vec{R}_j | \Phi_m \rangle = \langle \Phi_m | \sum_i e\vec{r}_i | \Phi_m \rangle + \sum_j Z_j |e| \vec{R}_j \quad ; \tag{28.170}$$

ici, les produits scalaires $\langle .|.|. \rangle$ sont des intégrations sur les seules coordonnées électroniques $\vec{r}_i$, les coordonnées nucléaires $\vec{R}_j$ étant les observables associées aux positions des noyaux. En conséquence, l'expression précédente se récrit toujours comme suit :

$$\vec{d} = \vec{D}_{\mathrm{e\,eff}}(\vec{R}_j) + \vec{D}_N(\vec{R}_j) \quad , \tag{28.171}$$

où $\vec{D}_{\mathrm{e\,eff}}(\vec{R}_j)$ est la contribution des électrons dont le mouvement rapide a été moyenné, et où $\vec{D}_N(\vec{R}_j)$ est la part des noyaux au moment dipolaire total.

[76] Il s'agit ici de la diffusion élastique résonnante de photons : tout comme pour un atome à deux niveaux, la susceptibilité est exaltée lorsque la fréquence du champ coïncide avec l'une des fréquences de Bohr du système. L'image classique est une vibration forcée du dipôle microscopique qui, en tant que telle, s'effectue à la fréquence de l'onde excitatrice.

Schématiquement, il s'agit de processus élémentaires du premier ordre vis-à-vis de la théorie des perturbations dépendant du temps. On sait que pour une onde quasi-monochromatique de pulsation ω et de durée T, la probabilité de transition fait apparaître une pseudo-fonction de Dirac $\delta_T(E_\mathrm{f} - E_\mathrm{i} - \hbar\omega)$ de largeur $\hbar/T$, traduisant la conservation de l'énergie au sens de Bohr : $E_\mathrm{f} = E_\mathrm{i} + \hbar\omega$, pour une transition en absorption (voir sous-section 24.3.2).

[77] La molécule étant neutre, $\vec{d}$ est invariant dans tout changement de l'origine du repère : c'est donc bien un attribut intrinsèque de la molécule.

Pour une molécule diatomique, le moment $\vec{d}$ est forcément porté par l'axe internucléaire $\overrightarrow{\mathrm{AB}}$ puisque, en tant qu'attribut de la molécule, il est invariant dans toute opération de symétrie :

$$\vec{d} = d(R)\frac{\vec{R}}{R} \ . \qquad (28.172)$$

Par ailleurs, si la molécule est homonucléaire (A = B, A_2 en notation chimique), le système est symétrique dans l'inversion par rapport au milieu de la liaison. Le seul vecteur invariant dans cette opération est le vecteur nul : dépourvue de moment dipolaire, une molécule de ce type est donc transparente dans l'infra-rouge et les micro-ondes[78]. Seules les molécules hétéronucléaires sont actives dans cette région spectrale ; pour une telle molécule, l'existence d'un moment dipolaire $\vec{d}$ permanent, plus ou moins grand, résulte physiquement du fait que, compte tenu des charges nucléaires différentes, les électrons sont plus attirés par un noyau que par l'autre (celui qui est le plus électronégatif), donnant une densité électronique dissymétrique et non invariante par réflexion (ou par parité).

Comme ce sont la vibration et la rotation de la molécule telle qu'elle existe qui sont pertinentes, la forme utile du moment dipolaire $\vec{d}$ résulte d'un développement de Taylor autour de R_0 de l'expression (28.171), étant entendu que $\vec{d}$ est le long de $\vec{R} \equiv \overrightarrow{\mathrm{AB}}$:

$$d(R) = d(R_0) + q_{\text{eff}}\,(R - R_0) + \ldots \ , \qquad (28.173)$$

où la constante q_{eff} est une charge électrique effective, en principe calculable à partir des fonctions d'onde électroniques (voir le passage de (28.170) à (28.171)). Si le champ est polarisé suivant Oz, le produit scalaire $-\vec{d}.\vec{\mathcal{E}}$ donne $d(R)\cos\theta$, θ étant l'angle entre l'axe de la molécule et Oz ; avec cette polarisation, les intensités des raies de vibration - rotation sont gouvernées[79] par le module carré de l'élément de matrice :

$$\langle F_{vLM}|d(R)\cos\theta|F_{v'L'M'}\rangle \ . \qquad (28.174)$$

Les variables étant séparées, l'expression (28.174) est le produit de trois intégrales. L'une, de 0 à 2π, porte sur l'angle azimutal ϕ, son intégrand est $\mathrm{e}^{\mathrm{i}(M'-M)\phi}$ et n'est non-nulle que si $M' = M$, d'où la règle de sélection[80] $\Delta M = 0$. Ceci étant fait, il reste le produit des deux intégrales :

$$\langle u_v|d(R)|u_{v'}\rangle\,\langle Y_{LM}|\cos\theta|Y_{L'M'}\rangle \ . \qquad (28.175)$$

Comme on l'a déjà vu à propos des transitions dipolaires électriques pour un atome (voir chapitre 24, sous-section (24.5.1)), le facteur angulaire n'est différent de zéro que si $|L - L'| = 1$, d'où une deuxième règle de sélection de type E1, $L' = L \pm 1$, soit en résumé :

$$\boxed{\Delta M = 0 \qquad \Delta L = \pm 1 \qquad (E1)} \qquad (28.176)$$

Enfin, le facteur radial de (28.174) est :

$$\langle \frac{1}{R}\,f_v|d(R_0) + q_{\text{eff}}(R - R_0)|\frac{1}{R}\,f_{v'}\rangle \ . \qquad (28.177)$$

[78] Voir plus loin la description qualitative de l'effet Raman.

[79] d'un point de vue purement mécanique. En temps utile, il faudra prendre en compte l'aspect *statistique* décrit par les probabilités de Boltzmannn de l'état de départ.

[80] Pour un champ électrique polarisé circulairement autour de l'axe de quantification Oz, la règle correspondante est $\Delta M = \pm 1$.

Les *bra - kets* désignent une intégration du genre $\int_0^{+\infty} R^2 f^*(R)\Omega(R)g(R)\,\mathrm{d}R$: compte tenu des deux facteurs $\frac{1}{R}$ dans f_v et $f_{v'}$, il apparaît strictement un élément de matrice avec des états propres d'oscillateur à une dimension :

$$\langle f_v|d(R_0) + q_{\mathrm{eff}}(R - R_0)|f_{v'}\rangle = d(R_0)\,\delta_{vv'} + q_{\mathrm{eff}}\,\langle f_v|(R - R_0)|f_{v'}\rangle \ , \qquad (28.178)$$

où $\langle f|\Omega|g\rangle = \int_0^{+\infty} f^*(R)\Omega(R)g(R)\,\mathrm{d}R \simeq \int_{-\infty}^{+\infty} f^*(x)\Omega(x)g(x)\,\mathrm{d}x$, puisque l'on ne doit considérer que des états vibrationnels peu excités. L'étude de l'oscillateur harmonique a montré que la coordonnée est proportionnelle à la somme $a + a^\dagger$ (Tome I, éq. (16.99)) : la matrice de la coordonnée sur les états propres du Hamiltonien n'a donc d'éléments de matrice non-nuls qu'entre deux états différant d'un quantum. D'où la troisième règle de sélection[81] :

$$\boxed{v' = v \quad \text{ou} \quad v' = v \pm 1} \qquad (28.179)$$

qui définit clairement deux classes de transitions, maintenant examinées successivement.

Spectre de rotation pure Les transitions $v' = v$ se font sans changement de l'état de vibration, et forment par définition le spectre de rotation pure. Compte tenu de (28.165), deux niveaux consécutifs de rotation pure sont séparés en énergie par :

$$B_m h[(L+1)(L+2) - L(L+1)] = 2B_m h(L+1) \qquad (L \geq 0)\ . \qquad (28.180)$$

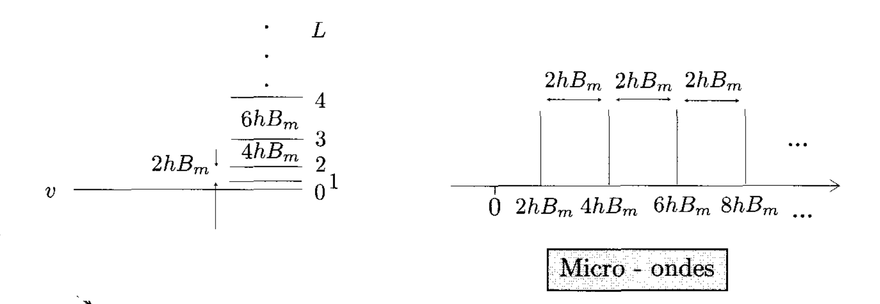


Figure 28.22: Schéma des transitions construisant le spectre de rotation pure ($v \to v$, $L \leftrightarrow L+1$).

En vertu de la règle $\Delta L = \pm 1$, le spectre de rotation pure est une série de raies impliquant les différences d'énergie $hB_m[(L+1)(L+2) - L(L+1)] = hB_m(2L+2)$ $(L \geq 0)$. Elles sont donc équidistantes, centrées en fréquence aux points $2B_m$, $4B_m$, $6B_m$, etc. (fig. 28.22). Outre la dépendance vis-à-vis de L et v des éléments de matrice, les intensités sont pondérées par les poids de Boltzmann[82] donnant les populations thermiques des états

[81] À la petite réserve près énoncée dans la note 75.

[82] À l'ambiante, $k_{\mathrm{B}}T \simeq 25\,\mathrm{meV}$. Cette énergie est nettement plus grande que l'échelle d'énergie donnant les différences d'énergie rotationnelle ($\nu_{\mathrm{rot}} \sim 5 \times 10^{10}$ Hz $\leftrightarrow 0{,}2\,\mathrm{meV}$). Il en résulte qu'à l'ambiante, les états rotationnels sont effectivement très distribués (voir problème 28.4.7).

de départ ; comme $g_L = 2L+1$ est la dégénérescence de chaque niveau L de rotation, ces probabilités sont de la forme $(2L+1)\mathrm{e}^{-AL(L+1)}$. La distribution de probabilité présente donc un maximum quand L varie, survenant pour $L^2 h B_m \sim k_\mathrm{B} T$.

Spectre de vibration-rotation Les transitions $v' = v \pm 1$ forment le spectre dit de vibration - rotation. Compte tenu des ordres de grandeurs, les degrés de liberté de rotation donnent en fait une structure fine aux raies de vibration, lesquelles se situent dans l'infra-rouge. On peut distinguer deux groupes de transitions :

1. $L \to L+1$, correspondant aux différences d'énergie :

$$\hbar\omega_m + B_m h[(L+1)(L+2) - L(L+1)] = \hbar\omega_m + 2B_m h\,(L+1)\ . \qquad (L \geq 0) \quad (28.181)$$

2. $L \to L-1$, correspondant aux différences d'énergie :

$$\hbar\omega_m + B_m h[(L-1)L - L(L+1)] = \hbar\omega_m - 2B_m h\,L\ . \qquad (L \geq 1) \quad (28.182)$$

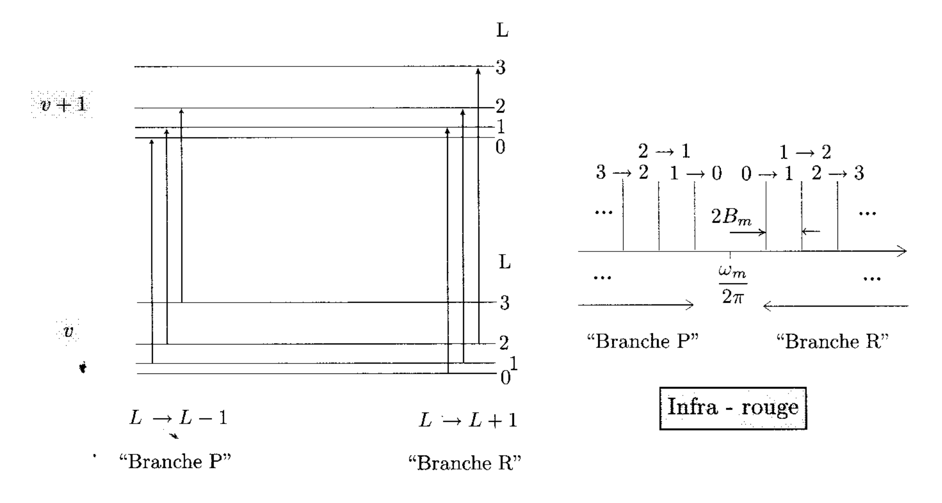


Figure 28.23: Schéma des transitions donnant le spectre de vibration-rotation ($v \to v+1$, $L \to L \pm 1$).

Au total, on obtient deux groupes de raies équidistantes, séparées de $2B_m$ en fréquence, symétriques de part et d'autre de $\nu_m = \frac{\omega_m}{2\pi}$. Finalement, tout se passe comme si la raie ($\Delta v = \pm 1$, $\Delta L = 0$) (qui n'existe pas), était décomposée par la structure fine due à la rotation (voir fig. 28.23). Tout comme pour la vibration pure, les intensités sont pondérées par des poids de Boltzmann ($\nu_\mathrm{vibr} \sim 5 \times 10^{12}$ Hz $\leftrightarrow$ 20 meV).

Ce qui précède rend compte, pour l'essentiel, des spectres de vibration - rotation. Il existe bien sûr diverses corrections à effectuer le cas échéant pour expliquer certaines

finesses. Le plus souvent, la plus importante d'entre elles s'obtient en développant le terme centrifuge :

$$\frac{\hbar^2 L(L+1)}{2\mu R^2} = \frac{\hbar^2 L(L+1)}{2\mu}\left[\frac{1}{R_0^2} - 2(R-R_0)\frac{1}{R_0^3} + 3(R-R_0)^2\frac{1}{R_0^4} + \ldots\right] \ . \quad (28.183)$$

Le terme linéaire décale légèrement la position d'équilibre de R_0 en $\tilde{R}_0$, cependant que la correction quadratique renormalise la pulsation ω_m en $\tilde{\omega}_m$. Enfin, il arrive aussi que l'on doive considérer les termes anharmoniques, c'est-à-dire introduire le terme en $(R-R_0)^3$ dans le développement de E_m autour de R_0 : alors les niveaux de vibration ne sont plus strictement équidistants (les écarts entre deux niveaux consécutifs diminuent quand on monte en énergie[83]). Le calcul convenable des diverses corrections exige du savoir-faire : il faut d'une part effectuer des calculs de perturbation aux ordres successifs, d'autre part comparer entre elles les importances relatives des différentes perturbations.

28.3.2 Effet Raman

On a vu qu'une molécule homonucléaire est inactive dans l'IR ou les micro-ondes : dans cette gamme spectrale, tous les éléments de matrice (dipolaires électriques) sont nuls. Pour obtenir le même type d'informations que celles obtenues sur les molécules AB par les moyens décrits ci-dessus, il convient donc de s'y prendre autrement, en exploitant un effet, dit *effet Raman*, dont le principe est donné ci-après en raisonnant semi-classiquement.

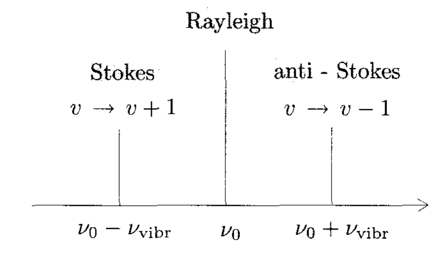


Figure 28.24: Schéma du spectre Raman de vibration ; quand la molécule gagne un *quantum* de vibration ($v \to v+1$), c'est au détriment du photon diffusé, qui a une fréquence plus *petite* que le photon incident, et inversement.

On connaît la diffusion Rayleigh, qui est la mise en vibration forcée des électrons d'un atome ou d'une molécule, à des fréquences petites devant la plus petite fréquence de résonance. D'une façon générale[84], lorsqu'un champ électromagnétique est envoyé sur

[83] Retenir l'image : plus le potentiel est raide, plus l'écart entre deux niveaux d'énergie consécutifs augmente vite, et inversement. Le spectre de l'oscillateur harmonique est uniforme (écarts constants, E_n varie comme n), les énergies du puits carré infini augmentent comme n^2 (écarts $E_{n+1} - E_n \propto n$), le puits infini étant visiblement plus raide que le puits harmonique (voir Tome I, fig. 16.1). À l'inverse, pour un potentiel ayant une limite à grande distance, l'écart entre deux niveaux successifs se réduit quand on monte en énergie (potentiel de Coulomb, de Morse, etc.).

[84] Voir Tome I, chapitre 6, et problème 6.4.5.

un tel système, et si sa fréquence[85] ν_0 ne coïncide pas avec une résonance, on obtient une diffusion élastique des photons dont la section efficace varie comme ν_0^4 à basse fréquence (diffusion Rayleigh) et tend vers une constante[86] σ_T à haute fréquence (diffusion Thomson, – diffusion élastique des rayons X, par exemple). Sous l'effet du champ excitateur, le petit dipôle induit est mis en oscillation forcée et rayonne à la fréquence ν_0 : la quantité importante est donc la *polarisabilité* de l'atome ou de la molécule. Cette polarisabilité dépend évidemment de la géométrie de la molécule ; si les noyaux sont fixes, une seule fréquence existe, c'est ν_0 et l'on obtient une simple diffusion élastique, généralement dans l'optique ou l'UV.

À l'inverse, si les noyaux vibrent à la fréquence $\nu_{\rm vibr}$, le dipole induit sera (lentement) modulé en amplitude à la fréquence $\nu_{\rm vibr}$, très petite devant ν_0. La combinaison des ces oscillations fait alors apparaître naturellement deux composantes supplémentaires $\nu_0 \pm \nu_{\rm vibr}$: outre la raie élastique Rayleigh (toujours présente, la molécule ne change pas d'état vibrationnel durant la diffusion du photon), deux raies symétriquement déplacées apparaissent (raies inélastiques), l'une en $\nu_0 + \nu_{\rm vibr}$ (raie dite[87] *anti-Stokes*), l'autre en $\nu_0 - \nu_{\rm vibr}$ (raie dite *Stokes*). Ces deux raies correspondent au cas où la diffusion modifie l'état vibrationnel de la molécule ($v \rightarrow v \pm 1$). La raie Stokes ($\nu < \nu_0$) est celle où la molécule gagne un *quantum* de vibration au détriment du photon, la raie anti-Stokes ($\nu > \nu_0$) celle où la molécule en perd un et où on récupère un photon diffusé de plus grande énergie (voir fig. 28.24).

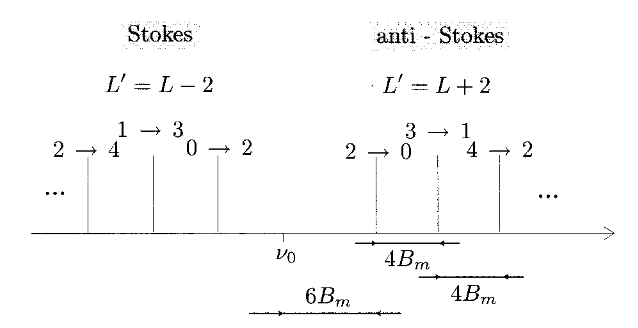


Figure 28.25: Schéma du spectre Raman de rotation ; quand la molécule passe de L à $L + 2$, la fréquence du photon diffusé est plus *petite* que ν_0, et inversement.

Le même effet vaut pour la rotation : une molécule qui tourne sur elle-même constitue un dipole oscillant (à une fréquence *double* : quand la molécule homonucléaire fait un demi-tour, le dipôle induit revient à l'identique). On attend donc aussi un effet Raman, et des raies Stokes et anti-Stokes, sur les spectres de rotation. Une différence existe cependant par rapport à la vibration : alors que l'oscillateur a une fréquence

[85] En pratique, pour la spectroscopie Raman, ν_0 est une fréquence optique.

[86] $\sigma_T = \frac{8\pi}{3} r_e^2 \simeq 6,7 \times 10^{-25}\,{\rm cm}^2$, où $r_e = \frac{e'^2}{mc^2} \simeq 2,82\,{\rm F}$ est le rayon classique de l'électron.

[87] Cette appellation est en honneur à George Gabriel Stokes (1819-1903), qui a apporté des contributions majeures à la dynamique des fluides (équation de Navier - Stokes), aux mathématiques (théorème de Stokes) et à la compréhension de la fluorescence. Le *Stokes shift* désigne généralement le décalage en fréquence (ou longueur d'onde) entre une excitation à ν_0 et une émission décalée à $\nu_0 \pm \Delta\nu$.

unique (dans l'approximation harmonique, le spectre d'énergie est une échelle à barreaux équidistants), il n'en va pas de même pour un rotateur. Il existe donc *a priori* une infinité de fréquences $\nu_0 \pm 2\nu_{\text{rot}}$ où $\nu_{\text{rot}} = 2B_m, 4\,B_m, \ldots$, donnant – *modulo* la règle de sélection[88] $\Delta L = 0, \pm 2$ – un grand nombre de raies de chaque sorte, Stokes ou anti-Stokes (fig. 28.25).

28.3.3 Molécules polyatomiques

Coordonnées normales

Pour une molécule ayant $N > 2$ noyaux, il existe $N_\text{V} = 3N - 6$ ou $3N - 5$ degrés de liberté de vibration, notés R_j. Le jeu des R_j conduit à une *déformation* de la molécule relativement à une configuration d'équilibre stable, prise comme référence :

$$E_m(R_j) \simeq E_m(R_{j0}) + \sum_{j,\,j'=1}^{N_\text{V}} \frac{1}{2}(R_j - R_{j0})(R_{j'} - R_{j'0}) \left(\frac{\partial^2 E_m}{\partial R_j \partial R_{j'}}\right)_0 , \qquad (28.184)$$

configuration autour de laquelle les noyaux effectuent de petites vibrations approximées par le développement harmonique généralisant (28.162). Comme la configuration est supposée stable, la fonction jouant le rôle d'une énergie potentielle harmonique est une forme quadratique définie positive que l'on peut diagonaliser en formant les bonnes combinaisons linéaires des R_j, appelées coordonnées normales et notées X_l. Au total, l'énergie électronique, dans l'approximation harmonique, peut se mettre sous forme diagonale :

$$E_m(R_j) = E_m(R_{j0}) + \sum_{l=1}^{N_\text{V}} \frac{1}{2} k_l\, X_l^2 \ . \qquad (28.185)$$

Toutes les constantes de raideur k_l sont strictement positives. En effectuant le même changement de variables dans les termes cinétiques constituant T_N, l'équation pour la fonction nucléaire prend la forme :

$$\Big(\sum_{l=1}^{N_\text{V}} H_{\text{harm},\,l} + H_\text{rot}\Big)\chi = E\,\chi \ . \qquad (28.186)$$

Le premier terme est une somme de N_V oscillateurs harmoniques à une dimension. Le second est le Hamiltonien du rotateur rigide constitué par la molécule dont les noyaux sont fixés aux positions d'équilibre, tout couplage entre vibration et rotation étant négligé, et incorpore les derniers degrés de liberté, la translation en bloc étant toujours mise de côté. Notons qu'une vibration normale implique en général tous les noyaux : c'est donc bien un mouvement *collectif*[89] de ces derniers.

[88] avec toujours $\Delta M = 0$.

[89] analogue à celui que l'on rencontrera pas la suite à propos des vibrations de réseau, dont la quantification produit les *phonons* (voir chapitre 31).

L'introduction des coordonnées normales peut aussi s'effectuer de façon plus formelle. On part de :

$$H_\mathrm{N} = \sum_{i=1}^{N} \frac{1}{2} M_i \vec{V}_i^{\,2} + E_m(Q_j) \ . \tag{28.187}$$

Posant $\vec{s}_j = \sqrt{M_j}\,(s_j - s_{j0})$, $\dot{\vec{s}}_j = \sqrt{M_j}\,\vec{V}_j$, le terme cinétique devient :

$$2T_\mathrm{N} = \sum_{i=1}^{N} \sum_{u=x,y,z} \dot{s_{ju}}^2 \equiv \sum_{\lambda=1}^{3N} \dot{s_\lambda}^2 \ , \tag{28.188}$$

d'où :

$$H_\mathrm{N} = \frac{1}{2} \sum_{\lambda=1}^{3N} \dot{s_\lambda}^2 + V(\{s_\lambda\}) \ , \qquad V(\{s_\lambda\}) \stackrel{\text{déf}}{=} E_m(Q_j(\{s_\lambda\})) \ . \tag{28.189}$$

En développant V autour de la position d'équilibre :

$$V(s_\lambda) = \sum_{\lambda,\mu=1}^{3N} \frac{1}{2} f_{\lambda\mu}\, s_\lambda s_\mu + \ldots \ , \tag{28.190}$$

on trouve finalement :

$$H_\mathrm{N} = \frac{1}{2} \sum_{\lambda=1}^{3N} \dot{s_\lambda}^2 + \sum_{\lambda,\mu=1}^{3N} \frac{1}{2} f_{\lambda\mu} s_\lambda s_\mu \ . \tag{28.191}$$

Maintenant, V est diagonalisé par une transformation orthogonale :

$$X_l = \sum_{\lambda=1}^{3N} O_{l\lambda} s_\lambda \ . \tag{28.192}$$

Cette transformation laisse invariant le carré de la norme ; il en résulte que le terme cinétique s'exprime comme suit :

$$T_\mathrm{N} = \frac{1}{2} \sum_{\lambda=1}^{3N} \dot{X_l}^2 \ , \tag{28.193}$$

cependant que, par définition, l'énergie potentielle devient :

$$V = \frac{1}{2} \sum_{\lambda,\mu=1}^{3N} \frac{1}{2} k_l X_l^2 \ . \tag{28.194}$$

On peut démontrer – c'est une évidence physique – 6 (5) constantes k_l sont nulles, associées à la rotation et à la translation en bloc. Autrement dit, l'équation caractéristique :

$$\begin{vmatrix} f_{11} - \Lambda & f_{12} & \cdots & f_{13N} \\ f_{21} & f_{22} - \Lambda & \cdots & f_{23N} \\ \vdots & \vdots & \ddots & \vdots \\ f_{3N1} & f_{3N2} & \cdots & f_{3N3N} - \Lambda \end{vmatrix} = 0 \tag{28.195}$$

possède 6 (5) fois la valeur propre $\Lambda = 0$.

La détermination des coordonnées normales est grandement facilitée par l'utilisation explicite de la symétrie moléculaire (rotations, réflexions, ...). En effet, soit S une opération de symétrie de la configuration d'équilibre ; par nature, effectuer cette opération ne change rien et, en particulier, ne modifie pas l'énergie potentielle. On a donc d'une part ($\Delta Q_j = Q_j - Q_{j0}$) :

$$E_m(SQ_j) - E_m(SQ_{j0}) = \sum_{j,\,j'} \frac{1}{2}(S\Delta Q_j)\,(S\Delta Q_{j'})\,H_{jj'} \ . \tag{28.196}$$

et, d'autre part :

$$E_m(SQ_j) - E_m(SQ_{j0}) = E_m(Q_j) - E_m(Q_{j0}) \ . \tag{28.197}$$

Par comparaison, on en déduit :

$$\sum_{j,\,j'} (S\Delta Q_j)\,(S\Delta Q_{j'})\,H_{jj'} = \sum_{j,\,j'} \Delta Q_j\,\Delta Q_{j'}\,H_{jj'} \ . \tag{28.198}$$

L'opération de symétrie est une transformation linéaire, représentée par une matrice orthogonale[90] :

$$S\Delta Q_j = \sum_k S_{jk}\Delta Q_k \qquad\qquad {}^{\mathrm{t}}S = S^{-1} \ . \tag{28.199}$$

D'où, successivement :

$$\sum_{j,\,j'}\sum_{k,\,k'} S_{jk}\Delta Q_k\, S_{j'k'}\Delta Q_{k'}\, H_{jj'} = \sum_{j,\,j'} \Delta Q_j\,\Delta Q_{j'}\,H_{jj'} \ , \tag{28.200}$$

$$\sum_{k,\,k'} \Big(\sum_{j,\,j'} S_{jk}H_{jj'}S_{j'k'}\Big)\Delta Q_k\,\Delta Q_{k'} = \sum_{j,\,j'} \Delta Q_j\,\Delta Q_{j'}\,H_{jj'} \ , \tag{28.201}$$

$$\sum_{k,\,k'} \left({}^{\mathrm{t}}S\,H\,S\right)_{kk'} \Delta Q_k\,\Delta Q_{k'} = \sum_{j,\,j'} \Delta Q_j\,\Delta Q_{j'}\,H_{jj'} \ . \tag{28.202}$$

Les indices étant muets, on en déduit :

$$\left({}^{\mathrm{t}}S\,H\,S\right)_{jj'} = H_{jj'} \quad \Longleftrightarrow \quad S^{-1}HS = H \quad \Longleftrightarrow \quad HS = SH \ ; \tag{28.203}$$

la dernière égalité n'exprime rien d'autre que, S étant une opération de symétrie, H et S commutent. Maintenant, soit X_l une certaine combinaison linéaire propre de S :

$$SX_l = \sigma_l X_l \ ; \tag{28.204}$$

comme $[H,\, S] = 0$, on a :

$$HSX_l = H(\sigma_l X_l) = \sigma_l HX_l = SHX_l \quad \Longleftrightarrow \quad S(HX_l) = \sigma_l(HX_l) \ . \tag{28.205}$$

[90] Puique toute opération de symétrie conserve les angles et les distances.

Ceci montre que si X_l est propre de S, alors HX_l est encore propre de S, avec la même valeur propre σ_l. Si σ_l n'est pas dégénérée, alors SX_l est proportionnel à X_l, c'est-à-dire que X_l est aussi propre de H :

$$HX_l = h_l X_l \ . \tag{28.206}$$

Autrement dit, X_l est l'une des directions principales de la forme quadratique donnant E_m dans l'approximation harmonique.

Si σ_l est dégénérée, deux fois par exemple, il existe deux combinaisons linéaires telles que :

$$SX_{l1} = \sigma_l X_{l1} \ , \qquad SX_{l2} = \sigma_l X_{l2} \ . \tag{28.207}$$

On a toujours :

$$S(HX_{l1}) = \sigma_l(HX_{l1}) \ , \qquad S(HX_{l2}) = \sigma_l(HX_{l2}) \ . \tag{28.208}$$

Comme HX_{li} $(i = 1,\, 2)$ est propre de S avec la valeur propre σ_l, chaque HX_{li} est une combinaison linéaire de X_{l1} et X_{l2} : on peut donc diagonaliser H dans ce sous-espace sans pour autant dédiagonaliser la restriction de S à ce sous-espace ; même en cas de dégénérescence, on peut rendre simultanément diagonales les deux matrices de H et S.

Ainsi, au lieu de procéder directement à la diagonalisation de la matrice de l'énergie potentielle, mieux vaut considérer d'emblée les opérations de symétrie et rechercher – au moyen des outils de la Théorie de la représentation linéaire des groupes – les combinaisons propres, ce qui assure une diagonalisation (ou au moins une prédiagonalisation) de l'énergie potentielle (28.184) (voir problème 28.4.8).

Rotation des molécules polyatomiques

Pour terminer, disons un mot de la rotation en bloc (de la configuration d'équilibre) – qui donne lieu à une quantification en raison du caractère périodique des variables angulaires et de la nécessité pour la fonction d'onde d'être monovaluée. En tant que corps solide, la molécule dans sa configuration d'équilibre possède un tenseur d'inertie I_{uv} $(u,\, v = x,\, y,\, z)$, dont les éléments sont définis comme suit :

$$I_{xx} = \sum_{\text{Noyaux}} M_i(y_i^2 + z_i^2) \ , \qquad I_{xy} = \sum_{\text{Noyaux}} M_i x_i y_i \ , \qquad \ldots \ , \tag{28.209}$$

où les coordonnées sont définies par rapport à un repère lié à la molécule et dont l'origine est le centre de masse G. Ce tenseur d'inertie est réel symétrique et peut donc être diagonalisé, ce qui conduit à trois axes principaux X, Y et Z. Sur ces directions, le tenseur d'inertie est représenté par la matrice diagonale :

$$\begin{bmatrix} I_{XX} & 0 & 0 \\ 0 & I_{YY} & 0 \\ 0 & 0 & I_{ZZ} \end{bmatrix} \ . \tag{28.210}$$

Si les trois moments d'inertie I_{UU} sont égaux entre eux, le rotateur est dit sphérique (ou isotrope). Si deux d'entre eux sont égaux, on parle de toupie symétrique, pour évoquer un objet ayant une symétrie de révolution cylindrique. Enfin, si les trois moments principaux sont tous différents, on a une toupie asymétrique.

L'énergie classique de rotation est purement cinétique (en l'absence de tout champ extérieur) et est de la forme :

$$\frac{1}{2}I_{XX}\omega_{XX}^2 + \frac{1}{2}I_{YY}\omega_{YY}^2 + \frac{1}{2}I_{ZZ}\omega_{ZZ}^2 \tag{28.211}$$

où les ω_{UU} sont les composantes du vecteur rotation $\vec{\Omega}$, donné par les vitesses angulaires de rotation autour des trois axes principaux. C'est bien parce que le tenseur d'inertie est diagonal vis-à-vis de ces derniers que l'énergie cinétique est une simple somme de trois termes indépendants les uns des autres ; pour chacune de ces directions principales, on a un rotateur rigide – qui se traite indépendamment des deux autres – comme on l'a vu à propos de la rotation d'une molécule diatomique.

28.4 Exercices et problèmes

28.4.1 Constantes du mouvement électronique pour une molécule diatomique

On se place dans la première étape de l'approximation de Born et Oppenheimer, où les noyaux sont fixes ; l'axe Oz est pris le long de la liaison chimique.

1. Soit $\vec{L}$ le moment cinétique électronique par rapport à un point quelconque de l'axe de la molécule. Quelle(s) composante(s) est (sont) une(des) constante(s) du mouvement ?
2. Soit M la symétrie miroir par rapport à un plan contenant l'axe de la molécule, associée à la transformation unitaire $\mathcal{M}$ dans l'espace des états. Combien vaut $[H, \mathcal{M}]$?
3. Qu'en est-il du commutateur $[\mathcal{M}, L_z]$?
4. Soit Ω un opérateur commutant avec H, et $|\psi\rangle$ un état propre de H d'énergie E. Que peut-on dire du vecteur $\Omega|\psi\rangle$?
5. Soit un vecteur propre commun à H et L_z, associé à la valeur $m = 0$ pour L_z, et à l'énergie E, supposée non-dégénérée. Est-ce aussi un vecteur propre de $\mathcal{M}$?
6. Soit les deux états $|E, \pm m\rangle$ associés à une même énergie et à deux valeurs opposées de L_z. La molécule est préparée dans l'état $|\Psi(t = 0)\rangle = \cos\alpha|E, m\rangle + \mathrm{e}^{\mathrm{i}\phi}|E, -m\rangle$; A étant une observable quelconque, la valeur moyenne $\langle\Psi(t)|A|\Psi(t)\rangle$ varie-t-elle au cours du temps ?

28.4.2 Méthode LCAO pour l'ion moléculaire H_2^+

On revient sur le traitement LCAO de l'ion H_2^+ (sous-section 28.2.2) ; les notations sont celles de la figure 28.4. La base d'orbitales atomiques est toujours réduite aux deux fonctions atomiques ϕ_i de type 1s centrées chacune sur un proton. On introduit les intégrales dites de recouvrement S, de Coulomb, C, et de résonance A :

$$S \stackrel{\text{déf}}{=} \langle\phi_1|\phi_2\rangle \ , \qquad C \stackrel{\text{déf}}{=} \langle\phi_1|\frac{e'^2}{r_2}|\phi_1\rangle \ , \qquad A \stackrel{\text{déf}}{=} \langle\phi_1|\frac{e'^2}{r_2}|\phi_2\rangle \ . \tag{28.212}$$

1. À l'aide de ces différentes quantités, écrire l'équation donnant les valeurs propres.
2. En raisonnant physiquement[91] montrer sans calcul que $C < \frac{e'^2}{R}$.
3. Calculer explicitement les intégrales S, C et A.
4. En déduire les deux valeurs propres $E_\pm(R)$, et $\Delta E_\pm \stackrel{\text{déf}}{=} E_\pm(R) - E_\pm(+\infty)$.
5. Trouver les formes approchées de ces différences pour $R \ll a_0$ et $R \gg a_0$; en déduire qu'il existe au moins un état stable.
6. À l'aide d'une machine, trouver la valeur R_0 du minimum de $E_-(R)$; combien vaut l'énergie de liaison ? Comparer aux valeurs expérimentales. Commenter.
7. Analyser les limites de l'énergie $E_{\text{el}}(R)$ quand $R \to 0$. Commenter.
8. Tracer en fonction de R la variation des énergies cinétique et potentielle.
9. Discuter qualitativement ce que donnerait le même traitement si on utilisait deux orbitales atomiques *strictement* localisées, c'est-à-dire telles qu'en tout point de l'espace l'une d'entre elles (au moins) est nulle.

28.4.3 Stabilité comparée des molécules He_2 et H_2

Expliquer qualitativement pourquoi la molécule He_2 est moins stable que H_2.

28.4.4 La molécule d'hydrogène selon Heitler et London

On reprend ici le traitement de Heitler et London, partant de la fonction d'onde (28.71).

1. Établir l'expression (28.80) des deux valeurs propres.

[91] On pourra invoquer le théorème de Gauss.

2. Soit r_1 et r_2 les distances d'un électron aux deux protons. En utilisant les coordonnées elliptiques définies en (28.41), et $\mathrm{d}^3r = \frac{1}{8}R^3(\xi^2 - \mu^2)\mathrm{d}\xi\mathrm{d}\mu\mathrm{d}\phi$, calculer les quantités $\tilde{\varepsilon}$ et $\tilde{v}$ définies en (28.76) et (28.77).

3. Il s'agit maintenant de calculer l'intégrale directe J définie en (28.78) en utilisant le développement (18.123).

   (a) En choisissant l'origine du repère sur l'un des protons, montrer que seul le terme $k = 0$ de ce développement donne une contribution non nulle.

   (b) Effectuer l'intégration sur l'une des deux variables radiales.

   (c) Achever le calcul en passant en coordonnées elliptiques.

   (d) À très grande distance, on attend $E_-(R) \simeq -\frac{C^{\text{ste}}}{R^\nu}$; quelle est la valeur de l'exposant ν ? Que donne le traitement de Heitler - London ?

28.4.5 Le polycétylène : limite $N\infty$ et analyse de la corrélation électronique

On modélise le polyacétylène comme une chaîne unidimensionnelle déformable constituée de N atomes à un électron de valence. Un électron localisé sur un site a la possibilité de sauter sur l'un des deux premiers voisins et interagit avec les autres électrons à la manière de Hubbard. La chaîne d'atomes est traitée dans une approximation classique où l'on oublie l'énergie cinétique des cœurs supposés d'inertie infinie, pour ne retenir que l'énergie élastique de déformation, prise dans l'approximation harmonique. Avec ces hypothèses, le Hamiltonien est :

$$H = -\sum_{n,s} t_{nn+1}(c_{n+1s}^\dagger c_{ns} + c_{ns}^\dagger c_{n+1s}) + \frac{U}{2}\sum_{n,s} c_{ns}^\dagger c_{ns} c_{n-s}^\dagger c_{n-s} + \frac{1}{2}K\sum_n (u_{n+1} - u_n)^2 \ , \tag{28.213}$$

soit $H \equiv T + V + H_\mathrm{L}$. u_n désigne l'écart à la position na qu'aurait le n^e ion si le réseau était rigide ; $s = \pm 1/2 \equiv \uparrow, \downarrow$. L'intégrale de saut t_{nn+1} dépend de la distance entre deux ions ; résultant essentiellement du recouvrement spatial entre deux fonctions localisées, on la modélise par une exponentielle :

$$t_{nn+1} = t_0 \mathrm{e}^{\alpha(u_n - u_{n+1})} \ . \tag{28.214}$$

α est une constante valant typiquement quelques Å^{-1}, t_0 est de l'ordre de l'électron-volt, K vaut quelques dizaines d'eV/Å^2. U est le paramètre de Hubbard, positif. Dans toute la suite, on pose $\gamma \stackrel{\text{déf}}{=} \frac{U}{2t_0}$, $\eta \stackrel{\text{déf}}{=} \frac{8\alpha^2 t_0}{K}$, $k_\mathrm{F} \stackrel{\text{déf}}{=} \frac{\pi}{2a}$, et on utilise des conditions cycliques.

Décrire qualitativement ce qui se passe dans les deux limites $\gamma \to 0$ et $\gamma \to +\infty$. L'étude qui suit permettra d'examiner la compétition entre la déformation de la chaîne et la corrélation électronique, et montrera une tendance à l'antiferromagnétisme lorsque les électrons sont en interaction forte.

Diagonalisation approchée de H

La dépendance de t_{nn+1} par rapport aux déplacements u_n rend le modèle insoluble[92]. On adopte une approximation de champ moyen revenant à linéariser le terme d'interaction :

$$\frac{U}{2}\sum_{n,s} c_{n\,s}^\dagger c_{ns} c_{n-s}^\dagger c_{n-s} \;\rightarrow\; U\sum_{ns}\langle N_{n-s}\rangle N_{ns} \qquad \Longleftrightarrow \qquad H \;\rightarrow\; H_{\text{MF}} \;. \tag{28.215}$$

1. On introduit les modes étendus $c_{ks}^\dagger \stackrel{\text{déf}}{=} N^{-1/2}\sum_n \mathrm{e}^{ikna}\, c_{ns}^\dagger$; compte tenu de la tendance à la dimérisation du réseau, on pose $u_n = (-1)^n u$. Exprimer l'opérateur cinétique T en fonction des opérateurs a et b définis comme :

$$a_{ks} \stackrel{\text{déf}}{=} c_{ks} \;, \qquad b_{ks} \stackrel{\text{déf}}{=} c_{k+2k_{\text{F}}s} \qquad (-k_{\text{F}} < k < +k_{\text{F}}) \;. \tag{28.216}$$

   En déduire T en fonction du champ ψ_{ks} à deux composantes $\begin{bmatrix} a_{ks} \\ b_{ks} \end{bmatrix}$ et des matrices de Pauli σ_u.

2. La tendance du système à développer des structures à $2k_{\text{F}}$ doit se retrouver également sur la densité électronique ; pour cette raison, on pose *a priori* :

$$\langle N_{ns}\rangle = \frac{1}{2} + (-1)^n\, 2s\,\delta \;, \tag{28.217}$$

   où δ est un paramètre à déterminer ultérieurement par une équation auto-cohérente. Compte tenu de ce choix, exprimer l'interaction V en fonction des a_{ks} et des b_{ks}, et mettre le Hamiltonien de champ moyen H_{MF} sous la forme :

$$H_{\text{MF}} = \sum_{k,s} \psi_{ks}^\dagger (h_0 \mathbf{1} - \vec{h}.\vec{\sigma})\psi_{ks} \;. \tag{28.218}$$

   Pour la suite, on pose $2s\gamma\delta + \mathrm{i}\sinh(2\alpha u)\sin ka = \rho\, \mathrm{e}^{\mathrm{i}\phi}$.

3. La diagonalisation de H_{MF} s'effectue par une transformation unitaire $\mathcal{S}$ telle que $\psi_{ks} = \mathcal{S}\Psi_{ks}$ où le champ Ψ_{ks} est construit sur des modes A et B à déterminer. Préciser en deux mots la structure de la transformation $\mathcal{S}$, puis la déterminer complètement. Trouver les ε_k tels que :

$$\frac{1}{2t_0} H_{\text{MF}} = \sum_{k,s} \Psi_{ks}^\dagger (h_0 \mathbf{1} - \varepsilon_k \sigma_z)\Psi_{ks} \;. \tag{28.219}$$

   Quelles sont les relations de commutation des opérateurs A_{ks} et B_{ks} ? Former l'état fondamental de la chaîne, $|\text{F}\rangle$.

4. Calculer la moyenne de N_{ns} dans l'état fondamental ; en déduire l'équation auto-cohérente à l'aide de la fonction elliptique de première espèce $\mathbf{K}(\lambda)$:

$$\mathbf{K}(\lambda) \stackrel{\text{déf}}{=} \int_0^{\pi/2} \frac{1}{\sqrt{1-\lambda\sin^2 x}}\,\mathrm{d}x \;, \qquad \mathbf{K}(\lambda \lesssim 1) \simeq \ln\frac{4}{1-\lambda} \;. \tag{28.220}$$

[92] Sur réseau rigide, ce problème a été résolu par Lieb et Wu en 1968 [241].

5. Trouver la valeur moyenne de l'énergie[93] dans l'état $|\mathrm{F}\rangle$; on posera :

$$\mathbf{E}(\lambda) \stackrel{\text{déf}}{=} \int_0^{\pi/2} \sqrt{1-\lambda\sin^2 x}\,\mathrm{d}x \ , \tag{28.222}$$

$\mathbf{E}(\lambda)$ étant la fonction elliptique de deuxième espèce.

6. Des résultats précédents, déduire l'énergie par particule $(2Nt_0)^{-1}E$ en fonction de $\Lambda \stackrel{\text{déf}}{=} (\cosh^2 2\alpha u + \gamma^2\delta^2)^{-1}$, γ, δ, η et $y \stackrel{\text{déf}}{=} 2\alpha u$.

Étude et comparaison des différentes solutions

1. Faire l'inventaire des différentes solutions et les décrire sommairement.
2. Soit ε_{D} l'énergie correspondant à $\delta = 0$. Préciser son comportement dans les deux limites[94] :
$$y \simeq 0 \ , \qquad y \gg 1 \ . \tag{28.223}$$
3. Écrire l'équation fixant y_0 correspondant au minimum de ε_{D}. Montrer qu'il existe une valeur η_c au-dessus de laquelle cette équation n'a plus de solution ; expliquer physiquement pourquoi il doit bien en être ainsi. Comment varie ε_{D} avec γ ?
4. Soit $\varepsilon_{\mathrm{SDW}}$ l'énergie de l'autre solution. A-t-elle un sens quel que soit y ? Préciser son point d'arrêt par rapport à l'abscisse y_{M} de son maximum.
5. Utiliser les résultats précédents pour situer les énergies ε_{D} et $\varepsilon_{\mathrm{SDW}}$ l'une par rapport à l'autre.
6. En déduire le diagramme de phase dans le plan (η, γ).
7. Établir l'expression du *gap* des excitations dans chacune des deux phases.

28.4.6 États liés du potentiel de Morse

Le potentiel de Morse est parfois utilisé pour représenter l'énergie électronique faisant office d'énergie potentielle pour le mouvement des cœurs d'une molécule diatomique. Il s'agit ici d'en trouver les états liés, servant de base à la description des vibrations ; il n'est plus alors nécessaire de faire l'approximation harmonique (petites oscillations). Soit, pour une particule de masse M, le potentiel de Morse à une dimension :

$$V(x) = V_0\big(\mathrm{e}^{-2\frac{x}{a}} - 2\mathrm{e}^{-\frac{x}{a}}\big) \qquad (V_0 > 0,\, a > 0) \tag{28.224}$$

[93] Pour le terme d'interaction, on utilisera l'égalité :

$$\langle \mathrm{F}|c_p^\dagger c_q^\dagger c_r c_s|\mathrm{F}\rangle = (\delta_{qr}\delta_{ps} - \delta_{qs}\delta_{pr})\langle \mathrm{F}|N_q|\mathrm{F}\rangle\langle \mathrm{F}|N_p|\mathrm{F}\rangle \ . \tag{28.221}$$

[94] $\mathbf{E}(\lambda \simeq 1) \simeq 1 + \frac{1}{4}(1-\lambda)\big(\ln\frac{16}{1-\lambda} - 1\big)$.

1. Tracer le graphe de $V(x)$. Préciser l'abscisse du minimum x_0 et la valeur $V(x_0)$.

2. On pose $V_0 = \frac{\hbar^2 k_0^2}{2M}$, $E = -\frac{\hbar^2 k^2}{2M}$ $(k > 0)$. Il s'agit maintenant de retrouver une équation différentielle connue, en effectuant des changements de variables et/ou de fonction en partant de l'équation aux valeurs propres pour les fonctions propres $\psi(x)$. Pour la clarté, on procède par étapes :

   (a) On pose $\psi(x) = \mathrm{e}^{-kaX} v(X)$, $X = \frac{x}{a}$. Écrire l'équation pour la fonction $v(X)$.

   (b) On effectue maintenant le changement de variable $\xi = \lambda \mathrm{e}^{-X}$, où λ est un paramètre à choisir judicieusement en temps utile, et $v(X) = \mathrm{e}^{-\xi/2} w(\xi)$. Écrire l'équation différentielle pour la fonction $w(\xi)$.

   (c) Montrer qu'il est possible de choisir λ de sorte que cette équation prenne la forme canonique de l'équation de Laplace (voir (19.137)).

3. La solution régulière est une certaine fonction hypergéométrique $\Phi(\alpha, \gamma, \xi)$; identifier les différents paramètres. En reprenant l'argument habituel, montrer que l'énergie des états liés est nécessairement quantifiée.

4. À quelle condition existe-t-il des états liés ? Commenter physiquement.

5. Exprimer le nombre d'états N_b liés en fonction du paramètre $k_0 a$.

6. Montrer que les énergie liées sont égales à :

$$E_n = -V_0 \Big(1 - \frac{n + \frac{1}{2}}{k_0 a}\Big)^2 , \qquad (28.225)$$

   où n est un entier positif ou nul.

7. Vérifier que, si $k_0 a \gg 1$, les premiers niveaux sont à peu près équidistants. Identifier la fréquence de fond de puits pour l'état fondamental.

8. À l'inverse, comment varie l'écart entre deux niveaux consécutifs très excités ?

28.4.7 Fonction de partition rotationnelle

Soit $Z(\beta)$ la fonction de partition pour les niveaux de rotation pure d'une molécule diatomique, dont le spectre est caractérisé par une certaine constante B_m (voir (28.164)).

1. Écrire l'expression de $Z(\beta)$.

2. En supposant $\beta h B_m \ll 1$, on peut approximer la somme discrète par l'intégrale correspondante ; en déduire l'expression approchée de $Z(\beta)$.

3. Tracer la distribution d'équilibre $P_L(\beta)$ des probabilités des diffférents états rotationels.

4. À l'aide d'une machine, comparer les distributions exactes et approchées pour plusieurs températures.

28.4.8 Modes normaux de vibration de X_3

Soit une molécule X_3, dont la configuration d'équilibre est un triangle équilatéral ; sachant que la molécule est infiniment rigide perpendiculairement au plan du triangle, il s'agit de construire les modes normaux (plans) vis-à-vis desquels la forme quadratique d'énergie potentielle harmonique est automatiquement diagonale. Comme il s'agit d'une forme quadratique symétrique réelle, la transformation diagonalisant celle-ci est *orthogonale.*

Le petit déplacement $\vec{u}_i$ de chaque atome autour de sa position d'équilibre est représenté par une double flèche (fig. 28.26), vecteur que l'on peut décomposer sur une base locale ($\vec{e}_i$, $\vec{e}_{i+1}$) comme indiqué sur la figure.

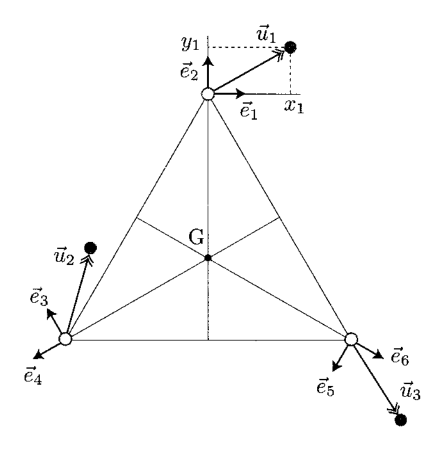


Figure 28.26: Bases locales permettant de repérer le déplacement de chaque atome par rapport à sa position d'équilibre.

Afin de construire des objets ayant des symétries données, on retrouvera certaines idées développées intuitivement dans le chapitre 22 à propos du groupe des permutations[95] $\mathcal{S}_N$. D'ailleurs, la question physique ici posée est l'occasion d'une initiation intuitive par l'exemple à la Théorie de la représentation linéaire des groupes ([50], [230]).

L'ensemble des opérations ponctuelles (à points fixes) de symétrie $\mathcal{G} \equiv \{S_i\}_{1\leq i\leq 6}$ contient ici l'identité (élément neutre) $S_1 \equiv \mathbf{1}$, deux rotations R_1 et R_2 autour d'un axe passant par G, et trois réflexions-miroirs M_1, M_2 et M_3 par rapport aux médiatrices des trois côtés. Cet ensemble peut être doté d'une loi de composition interne évidente : le produit S_iS_j consiste à appliquer d'abord S_j, puis S_i à la configuration d'équilibre. Vérifier qu'il s'agit bien d'un groupe ; pour simplifier, on notera $R_1 \equiv R$ et $R_2 = R^2$.

1. En énumérant les degrés de liberté, et en éliminant ceux qui ne donnent pas d'énergie élastique, trouver *a priori* le nombre de modes normaux de vibration.

[95]Il existe un lien important entre tout groupe de cardinal (ordre) fini, N, et le groupe symétrique $\mathcal{S}_N$, précisé par le théorème de Cayley : *"Tout groupe d'ordre N est isomorphe à un sous-groupe de $\mathcal{S}_N$"*.

2. Dresser la table de multiplication du groupe (vérifier au passage le lemme de réarrangement, invoqué au chapitre 22, section 22.3).

3. Ce groupe est-il abélien ?

4. Intuitivement, la classe d'équivalence $\mathcal{C}(S_i)$ d'une opération S_i donnée est formée des opérations qui ressemblent à S_i ; la définition précise est

$$\mathcal{C}(S_i) \stackrel{\text{déf}}{=} \{S_j S_i S_j^{-1}\}_{1 \le j \le 6} \ ; \tag{28.226}$$

former les classes d'équivalence de $\mathcal{G}$.

5. Tout écart par rapport à la configuration déquilibre est complètement défini par la donnée de 6 nombres x_1, y_1, ..., y_6 ; les 6 vecteurs des trois bases locales sont ainsi la base d'un espace vectoriel $\mathcal{E}$ sur $\mathbb{R}$, isomorphe à $\mathbb{R}^6$, dont chaque point représente une "position" de la géométrie déformée considérée ; c'est en quelque sorte l'espace des états de la molécule. Pour distinguer la base de $\mathcal{E}$ des bases locales, on pourra introduire l'isomorphisme $\vec{e}_i \leftrightarrow |e_i\rangle$. Le cas échéant, on utilisera le produit scalaire $(|e_i\rangle, |e_j\rangle) \equiv \langle e_i|e_j\rangle = \delta_{ij}$.

   Chaque opération S_i du groupe est ainsi *représentée* par une matrice carrée[96] $\mathcal{S}_i$ de dimension 6. De toute évidence, si $S_iS_j = S_k$, alors $\mathcal{S}_i\mathcal{S}_j = \mathcal{S}_k$: le produit des matrices représente fidèlement le produit des éléments du groupe. C'est pourquoi on dit de $\mathcal{E}$ qu'il forme l'*espace support* pour une représentation (linéaire) du groupe $\mathcal{G}$.

   Écrire quelques matrices $\mathcal{S}_i$ et vérifier l'assertion précédente.

6. Tout comme dans le chapitre 22 à propos des permutations, il s'agit d'engendrer systématiquement des vecteurs $|\Psi\rangle \in \mathcal{E}$ ayant une certaine symétrie ; pour les mêmes raisons, il est particulièrement utile de définir des projecteurs qui, agissant sur un vecteur quelconque, engendrent des vecteurs ayant justement la bonne symétrie, ou un vecteur nul si le vecteur d'épreuve est déjà d'une symétrie opposée.

   Soit par exemple le projecteur $P_{\text{S}} \stackrel{\text{déf}}{=} C\sum_{i=1}^{6} \mathcal{S}_i$, où C est une constante de normalisation assurant que $P_{\text{S}}^2 = P_{\text{S}}$. Combien vaut C ?

7. $|\psi\rangle$ désignant un vecteur quelconque de $\mathcal{E}$, on forme $|\Psi_{\text{S}}\rangle \stackrel{\text{déf}}{=} P_{\text{S}}|\psi\rangle$. À quoi sont égaux les six vecteurs $\mathcal{S}_i|\Psi_{\text{S}}\rangle$? La réponse justifie que l'on dise du vecteur $|\Psi_{\text{S}}\rangle$ qu'il se transforme suivant la représentation symétrique de $\mathcal{G}$, notée Γ^{S}.

8. Soit maintenant le projecteur $P_{\text{A}} \stackrel{\text{déf}}{=} C'(\mathbf{1} - \mathcal{M}_1 - \mathcal{M}_2 - \mathcal{M}_3 + \mathcal{R} + \mathcal{R}^2)$. Trouver C' de sorte que $P_{\text{A}}^2 = P_{\text{A}}$. Ayant formé $|\Psi_{\text{A}}\rangle \stackrel{\text{déf}}{=} P_{\text{A}}|\psi\rangle$, calculer les vecteurs $\mathcal{S}_i|\Psi_{\text{A}}\rangle$. Que constate-t-on ? Pour cette raison, $|\Psi_{\text{A}}\rangle$ est dit se transformer suivant la représentation antisymétrique de $\mathcal{G}$, notée Γ^{A}.

9. Que valent les produits $P_{\text{S}}P_{\text{A}}$ et $P_{\text{A}}P_{\text{S}}$? Qu'en déduit-on pour le produit scalaire $\langle\Psi_{\text{A}}|\Psi_{\text{S}}\rangle$?

[96] On adopte la même convention d'écriture que dans le chapitre 17 : une capitale en style romain S pour une opération dans l'espace physique, une capitale en calligraphie $\mathcal{S}$ pour son image dans l'espace des états $\mathcal{E}$.

10. On observe que les deux projecteurs ainsi formés sont de la forme[97] $\frac{1}{6}\sum_{i=1}^{6}\chi_i^{(\mathrm{I})}\mathcal{S}_i$, où les $\chi_i^{(\mathrm{I})}$ sont des nombres (ici réels), appelés *caractères* de la représentation Γ^{I}. Il existe une autre représentation, notée Γ^{E}, dont le projecteur correspondant P_{E} se forme de la même façon, en posant $P_{\mathrm{E}} = C''\sum_{i=1}^{6}\chi_i^{(\mathrm{E})}\mathcal{S}_i$. Trouver la constante C'' assurant que P_{E} est idempotent ; déterminer les $\chi_i^{(\mathrm{E})}$ assurant que les vecteurs $P_{\mathrm{S}}|\psi\rangle$, $P_{\mathrm{A}}|\psi'\rangle$ et $P_{\mathrm{E}}|\psi''\rangle$ sont orthogonaux quels que soient $|\psi\rangle$, $|\psi'\rangle$ et $|\psi''\rangle$.

    On remarquera que tous les éléments d'une même classe d'équivalence ont les mêmes caractères pour une représentation donnée ; ce résultat est général.

11. Une fois trouvés les projecteurs de symétrie (parfois appelés *opérateurs métriques*), on peut engendrer systématiquement les vecteurs de symétrie adaptée en les faisant agir sur n'importe quel vecteur de $\mathcal{E}$. Partant de $|e_2\rangle$, par exemple, trouver un vecteur de type $|\Psi_{\mathrm{S}}\rangle$; à quoi est égal $P_{\mathrm{S}}|e_1\rangle$? (commenter)

12. Recommencer ces opérations pour la représentation Γ^{A}.

13. De même, former $|\Psi_{\mathrm{E1}}\rangle \stackrel{\text{déf}}{=} P_{\mathrm{E}}|e_1\rangle$ et $|\Psi_{\mathrm{E2}}\rangle \stackrel{\text{déf}}{=} P_{\mathrm{E}}|e_2\rangle$. Ces deux vecteurs sont-ils linéairement indépendants ? En déduire par orthogonalité deux autres vecteurs $|\Psi'_{\mathrm{E1}}\rangle$ et $|\Psi'_{\mathrm{E2}}\rangle$ se transformant également suivant Γ^{E}.

14. Revenant à l'espace physique, former les combinaisons linéaires des $\vec{e}_i$ associés aux six vecteurs $|\Psi_{\mathrm{I=A,\,S}}\rangle$, $|\Psi_{\mathrm{E}\alpha=1,\,2}\rangle$ et $|\Psi'_{\mathrm{E}\alpha=1,\,2}\rangle$ trouvés ci-dessus.

15. Montrer que deux parmi ces vecteurs correspondent à des translations et un autre à une rotation. Qu'en déduit-on pour l'énergie potentielle ?

16. Les trois autres vecteurs correspondent réellement à des déformations du squelette moléculaire d'équilibre : dessiner les déplacements collectifs (modes normaux) correspondants.

17. L'un des modes normaux est de type Γ^{A}, les deux autres du type Γ^{E} ; comment se transforment ces derniers par $\mathcal{M}_1$? Qu'en déduit-on pour les deux valeurs propres associées de la forme quadratique diagonale représentant l'énergie potentielle harmonique exprimée avec les modes normaux ainsi déterminés ?

[97] Pour le groupe $\mathcal{S}_N$ rencontré dans le chapitre 22, les $\chi^{(\mathrm{S})}$ valent tous $+1$ (bosons), les $\chi^{(\mathrm{A})}$ sont les signatures des permutations (fermions).

Chapitre 29

Matière condensée ordonnée

Après une présentation des différents types de cristaux, la théorie élémentaire de l'énergie de cohésion est exposée pour les solides moléculaires et ioniques, et pour les métaux. On discute ensuite les bases physiques de la diffusion (rayons X et neutrons). S'agissant d'un ensemble de diffuseurs organisés dans l'espace, la condition de von Laue est démontrée ; elle permet de comprendre l'origine de l'hypersélectivité de la diffusion coopérative par un réseau, et constitue la signature de l'ordre à longue portée d'un cristal parfait.

29.1 Préliminaires

En raison de la complexité technique imposée par la Mécanique quantique, la description physique des systèmes à plusieurs corps exige de prendre quelque distance avec la rigueur formelle. On en a vu déjà quelques exemples ; le traitement d'un atome complexe peut être entrepris en distinguant les électrons internes de ceux de la périphérie de l'atome, le bon critère étant leurs énergies de liaison respectives, vite très différentes quand le numéro atomique croît – une différence marquée déjà réalisée pour des éléments légers puisque, pour le carbone, les électrons 1s ont une énergie de liaison de plusieurs centaines d'eV, alors que ceux responsables de la chimie du carbone et des transitions optiques impliquent des énergies de l'ordre de quelques eV, tout au plus.

Une dichotomie analogue a été faite pour les molécules, sur les mêmes bases, en considérant les bonnes échelles d'énergie, et a permis de distinguer les électrons engagés dans les liaisons et ceux très peu sensibles aux divers réarrangements qui se produisent inévitablement lorsque la molécule se forme. À cette occasion a été introduite la notion

qualitative de *cœurs* pour désigner ces fantômes des atomes libres, une fois la molécule constituée.

Ces considérations gardent évidemment toute leur pertinence lorsqu'il s'agit de matière *condensée* de basse énergie[1], qu'il s'agisse de liquides ou de solides. Lors de la transition de phase[2] faisant passer un système de l'état gazeux à l'état condensé, les constituants libres (atomes, molécules) conservent une part de leur intégrité, variable suivant leur nature, laquelle détermine pour une grande part celle du système une fois condensé. Les électrons complémentaires sont traditionnellement appelés *électrons de valence* en ce qui concerne ceux qui sont principalement responsables de la cohésion du solide, et, s'il en reste encore, *électrons de conduction* puisque, relativement libres et mobiles, ils sont *a priori* aptes à participer à la conduction électrique.

Ces faits avérés sur le plan expérimental autorisent tout naturellement une *classification* des solides, dont les différentes catégories sont caractérisées par un ensemble de propriétés à la fois bien définies et bien marquées même si, dès qu'une classification est établie, on peut rencontrer des cas marginaux par la proximité qu'ils entretiennent avec la frontière entre deux catégories.

À ce stade, quelques ordres de grandeur permettent de fixer les idées, et de prendre conscience quantitativement de ce qui sépare d'emblée un gaz d'un liquide et d'un solide. Il suffit pour cela d'estimer la distance moyenne entre constituants microscopiques d'une phase à l'autre. Dans les conditions normales, une mole de gaz parfait occupe environ 22,4 l ; avec le nombre d'Avogadro, on en déduit la densité n de particules et on trouve qu'elle est de l'ordre de $3 \times 10^{19}\,\mathrm{cm}^{-3}$. La quantité $d \stackrel{\text{déf}}{=} n^{-1/3}$ est une indication de la distance moyenne entre deux particules, et vaut $d_{\text{gaz}} \simeq 30\,\text{Å}$ pour un tel gaz : le milieu gazeux est très dilué, la distance entre deux objets étant très grande devant la taille de chacun d'entre eux. La masse volumique du gaz s'en déduit : pour l'air (4/5 d'oxygène, 1/5 d'azote en composition molaire, approximativement), on trouve $\rho_{\text{air}} \simeq 1,3\,\mathrm{g/l}$.

Pour un liquide, la masse volumique est de l'ordre du kilogramme par litre, soit $10^3\,\mathrm{kg/m^3}$. Une mole d'eau a pour masse 18 g, ce qui donne $n_{\text{liquides}} \simeq 3 \times 10^{22}\ \mathrm{cm}^{-3}$ et une distance moyenne entre atomes (molécules) $d_{\text{liquides}} \simeq 3\,\text{Å}$: la distance entre deux particules est du même ordre de grandeur que la taille de chacune d'entre elles ; un liquide est donc un milieu dense – mais fluide ! –, dont on peut se faire une image en disant que les particules y *roulent* les unes sur les autres. Enfin, la densité des solides est nettement plus élevée : pour le fer, par exemple ($M_{\text{Fe}} = 56\,\mathrm{g}$), la densité est $7{,}8\,\mathrm{g/cm^3}$, qui donne $n_{\text{solides}} \sim 10^{23}\,\mathrm{cm}^{-3}$ d'où $d_{\text{solides}} \sim 2\,\text{Å}$. Bien sûr, cette distance est d'ordre atomique, et est un peu plus petite que dans un liquide.

[1]La matière nucléaire est une matière pour le moins *condensée* ! On ne considère ici que des phénomènes impliquant des énergies situées dans la gamme de l'électron-volt, au plus.

[2]Pourquoi un corps, dans des conditions données, est-il plutôt dans un état (phase) que dans un(e) autre ? La réponse est fournie par la théorie des transitions de phase et plus particulièrement par celle des phénomènes critiques. Il s'agit encore d'une discipline en soi, qui ne sera pas abordée ici, d'autant plus que les phénomènes critiques connus en Thermodynamique sont gouvernées par une physique *à grande échelle* où les détails microscopiques ne sont pas pertinents et où l'*universalité* est la règle. Il existe cependant des transitions de phase *quantiques* à température nulle où, *mutatis mutandis*, les fluctuations quantiques jouent en un sens le rôle des fluctuations thermiques à température finie.

Ces estimations montrent que les solides et les liquides sont à considérer *a priori* sur le même plan[3], ce d'autant plus que force est d'admettre qu'il existe des liquides *solides*, communément appelés *verres*. Ces matériaux tiennent des solides par leur viscosité (finie pour un liquide, infinie pour un verre) mais sont similaires aux liquides par l'organisation microscopique de la matière, caractérisée par un ordre à courte portée (une dizaine d'Å) en tout point comparable à celui d'un liquide ordinaire. L'absence d'ordre sur une longue distance justifie que l'on dise des verres qu'ils sont *amorphes*, et explique que leur comportement physique, très complexe[4], est fort difficile à décrire théoriquement.

Dans la suite, on ne parlera ni de verres, ni d'amorphes : tous les solides examinés seront supposés *cristallins* ; pour un cristal (parfait), l'organisation spatiale microscopique de la matière obéit à des règles géométriques simples conduisant à la notion de réseau de Bravais et de motif. Les notions indispensables seront brièvement exposées en temps utile (voir sous-section 29.3.1).

Une première grande classification des solides se crée spontanément sur la considération des propriétés électriques permettant de séparer nettement conducteurs et isolants, sur la base de leur résistivité électrique dont la Mécanique quantique fournit une explication cohérente (voir chapitre 30, section 30.5). Il s'avère en outre très utile d'établir une sous-classification pour les isolants, fondée, elle, précisément sur les caractéristiques de la densité électronique. On distingue ainsi principalement :

- Les solides covalents, où l'on retrouve les caractéristiques *directionnelles* des liaisons chimiques. Le prototype de cette espèce de solide est le diamant, constitué d'atomes de carbone répartis dans l'espace suivant deux réseaux cubiques à faces centrées entremêlés, résultat de la valence 4 du carbone (structure tétraédrique[5]). D'autres exemples : le silicium et le germanium, ou même le carbone sous sa forme graphite[6] – qui ont pourtant des propriétés électriques très différentes du diamant ! Alors que le diamant est un isolant remarquable, on sait bien que Ge et Si – même purs – sont des semi-conducteurs[7] ; quant au graphite, il est difficile à classer :

[3]Il vaut la peine de comparer les valeurs numériques de la matière de basse énergie, à celles des milieux nucléaires. Le proton et le neutron ont à peu près la même masse M, soit $M \simeq 1,6 \times 10^{-25}$ g ; le rayon du noyau est de l'ordre de $1\,\mathrm{F} = 10^{-15}$ m (voir Tome I, chapitre 8), d'où une masse volumique $\rho_{\mathrm{noyau}} \sim 1,6 \times 10^{14}\,\mathrm{g/cm^3}$ – un dé à coudre de matière nucléaire “pèse” cent millions de tonnes. La densité de nucléons est elle aussi faramineuse : un nucléon par Fermi cube correspond à 10^{38} nucléons par cm^3...

[4]La physique des verres est encore très imparfaitement comprise. En particulier, les phénomènes de relaxation au sein des verres (à la suite, par exemple d'une contrainte mécanique transitoire) sont très difficiles à décrire quantitativement, notamment parce qu'ils relèvent de mécanismes *non-linéaires*, que l'on retrouve d'ailleurs dans d'autres systèmes *a priori* fort différents (frottement solide, tas de sable, tremblements de terre...). Un scénario typique est celui d'une accumulation de micro-contraintes qui tiennent le coup jusqu'à ce qu'une *avalanche* finisse par se déclencher (dans une avalanche de neige, il y a comme du tas de sable (!?) et du frottement solide (!?))...

[5]Comme signalé dans le chapitre 28, sous-section 28.2.4, c'est l'hybridation sp^3 qui est pertinente.

[6]expliquée par l'hybridation sp^2 produisant les angles de 120°.

[7]La différence tient bien sûr à la structure de bande et aux valeurs *numériques* des *gaps* d'énergie (à l'ambiante : 0,67 eV et 1,12 eV pour Ge et Si respectivement, 5,5 eV pour le diamant. Ces valeurs ne sont pas d'un ordre de grandeur différent, mais rentrent dans des exponentielles, d'où les ordres de grandeurs

sa conductivité augmente avec la température jusqu'à 80 K (comportement typique d'un semi-conducteur intrinsèque), puis décroît entre 300 K et 3 300 K. Les fondements théoriques de la différence entre isolants et conducteurs seront exposés plus loin (chapitre 30, section 30.5).

- Les solides moléculaires, qui sont essentiellement une juxtaposition d'atomes ou de molécules très peu perturbés par rapport à leur état libre en phase diluée. Il y a peu de densité électronique entre les atomes. Les prototypes sont les solides de gaz rares (Néon, Argon, Krypton, Xénon[8]) et les solides de certaines molécules aromatiques (anthracène, par exemple).

- Les solides ioniques, par exemple le sel ordinaire (NaCl). Ils sont formés à partir de couples d'atomes ; l'un des deux est de type métallique (il perd un ou des électrons lors de la formation du solide, c'est ce que fait Na), alors que l'autre est très avide d'électrons (très électronégatif, *antimétallique*, c'est le cas de Cl). Les électrons échangés restent très bien localisés près des centres, comme dans un cristal moléculaire, ce qui produit une structure formée de paires d'ions ; mais la réorganisation est importante par rapport à la structure des deux partenaires non-appariés, puisque l'un a perdu des électrons alors que l'autre en a gagné. Au total, les électrons restant bien localisés, l'interaction entre ions n'est pas écrantée et c'est l'interaction de Coulomb nue qui est à l'œuvre – bien évidemment complétée par une répulsion de cœur dur qui prévient le *collapse*.

En tout état de cause, et comme pour une molécule, ce sont bien sûr les électrons plus ou moins partagés entre les *briques* du solide qui constituent la colle, le ciment, de celui-ci. Selon le cas, on parlera de liaison métallique, liaison covalente ou liaison ionique. En ce qui concerne les solides moléculaires, il s'agit plutôt d'une pseudo-liaison chimique (fragile) résultant d'une interaction du type van der Waals, interprétable comme l'a vu (chapitre 25, section 25.4) en terme d'échanges de photons (virtuels) entre deux objets (atomes, molécules) globalement neutres. De façon plus élémentaire, il s'agit d'une interaction dipôle induit - dipôle induit : la présence du premier partenaire (et des charges électriques qu'il renferme) induit un moment dipolaire électrique d_2 au sein du deuxième partenaire, *via* sa polarisabilité α :

$$d_2 \sim \alpha\,\mathcal{E} = \alpha\,\frac{d_1}{R^3} \quad . \tag{29.1}$$

Comme l'interaction entre deux dipôles est du genre $\frac{d_1 d_2}{R^3}$, on en déduit une énergie d'interaction attractive variant comme $\frac{\alpha d_1}{R^3}\frac{d_1}{R^3} \propto R^{-6}$. Par $\vec{F} = -\vec{\nabla}\,V$, il en résulte que les forces de van der Waals décroissent comme R^{-7} – aux distances grandes par rapport à l'échelle atomique, mais pas trop grandes[9].

différents pour la conductivité – sans compter avec les valeurs prises par la densité électronique.

Au voisinage de l'ambiante, le *gap* varie linéairement en température. Ceci provient de la dilatation du réseau, qui induit une expansion du potentiel périodique.

[8]L'hélium fait bande à part : en vertu de sa petite masse, les effets quantiques ne peuvent jamais être ignorés et toute tentative de description classique est vouée à l'échec.

[9]Quand les photons (virtuels), vecteurs de l'interaction électromagnétique, mettent trop longtemps pour aller d'un atome à l'autre, il apparaît des corrections donnant finalement une décroissance plus rapide avec la distance (par exemple : $V \propto R^{-7}$, force en R^{-8} [200]).

La physique de la matière condensée est un monde en soi, par la variété et la richesse des phénomènes qui s'y produisent. La Mécanique quantique y a trouvé un champ d'application dont les frontières n'ont cessé de reculer, mais dont les chapitres suivants ne pourront donner qu'une vision très réduite. Parallèlement, les avancées en Physique atomique (les réseaux optiques notamment, les atomes froids, etc.) ont permis de transposer des phénomènes parfois surprenants sur d'autres échelles physiques (voir par exemple le problème 30.7.13 sur les oscillations de Bloch). Avant d'entreprendre une exploration forcément très limitée, il est, d'un point de vue physique, indispensable de comprendre pourquoi un solide (cristallisé) existe (*tient debout*) et de dire un mot des méthodes expérimentales qui permettent d'en connaître la structure. C'est le sujet des deux sections suivantes, où l'on donne d'abord une théorie élémentaire de la cohésion d'un solide, avant de faire un panorama des principes et des techniques de diffraction.

29.2 Énergie de cohésion des solides ordonnés

29.2.1 Présentation

L'énergie de cohésion d'un solide est l'énergie qu'il faut fournir pour le séparer en ses constituants à l'état libre. Une telle définition est clairement ambiguë, car tout dépend des fragments que l'on définit. D'une façon générale, on choisit les éléments libres au vu de considérations physiques mais tout choix reste forcément un peu conventionnel ; les distinctions qui en résultent sont de peu d'intérêt. Par exemple, s'agissant de l'azote solide, on pourrait décider de décomposer le solide en molécules N_2 plutôt qu'en atomes N, mais peu importe au fond, puisque l'on connaît l'énergie de dissociation de la molécule N_2.

Afin de rester au niveau le plus simple, la discussion ci-dessous est conduite avec l'hypothèse de cœurs *fixes*, situés aux nœuds d'un réseau géométrique. Il s'agit bien sûr d'une approximation, à la fois qualitative et quantitative : une particule ne peut être localisée strictement et est soumise aux inégalités de Heisenberg. Pour une particule de masse M, à une incertitude en position Δx correspond une énergie[10] $\Delta E \sim \Delta p^2/M$ soit $\Delta E \sim \hbar^2/(M\Delta x^2)$ et qui devrait être prise en compte dans le calcul de l'énergie de cohésion. Avec $\Delta x \sim 0,5\,\text{Å}$, et pour un atome de carbone, on trouve $\Delta E \sim 0,05\,\text{eV}$, ce qui est loin d'être négligeable devant l'énergie par atome que l'on trouvera dans l'hypothèse où le réseau est rigide ; clairement, le traitement des cœurs comme des objets ponctuels classiques sera d'autant meilleur que leur masse est plus élevée. Finalement, ce cadre est l'équivalent pour un solide de la première étape de l'approximation de Born et Oppenheimer, où les noyaux (cœurs) sont considérés comme totalement inertes du point de vue mécanique (masses réputées infinies). Dans les différentes situations physiques rencontrées dans la suite, l'énergie E, dite de cohésion, sera le pendant de l'énergie électronique d'une molécule ; tout comme cette dernière, paramétrée par les coordonnées nucléaires et dont le minimum vis-à-vis de ces dernières définit le squelette moléculaire d'équilibre, l'énergie de cohésion du cristal dépendra de grandeurs auxiliaires (paramètres

[10] appelée énergie de point zéro quand il s'agit de vibrations.

de la maille cristalline, longueur caractéristique liée à la densité du gaz d'électrons,...). L'énergie de cohésion physique retenue sera définie par le minimum de E par rapport à l'ensemble de ces paramètres, analogue de l'énergie de liaison pour une molécule diatomique dont le minimum définit la longueur de la liaison chimique.

Il est possible de construire une théorie élémentaire de la cohésion pour les solides moléculaires et ioniques, n'utilisant que les idées classiques mais conduisant à des résultats tout à fait acceptables. En revanche, rien de tel n'existe pour les solides covalents – sous-jacente se tient la théorie de la liaison chimique, dont on a entrevu la subtile complexité au chapitre 28. Quant aux métaux, la situation est nettement plus difficile à décrire proprement : on ne peut pas ignorer les électrons et leurs interactions, même si elles sont souvent assez fortement écrantées, en raison notamment de la grande densité électronique.

29.2.2 Cohésion des solides moléculaires

Le prototype du solide moléculaire est le cristal de gaz rare. Excluant l'hélium – si léger que les effets quantiques de point zéro ne peuvent être ignorés – on se place dans l'approximation où les atomes constitutifs sont traités comme des objets ponctuels classiques. Par définition d'un cristal moléculaire, ses atomes sont très peu modifiés par rapport à leur état libre lors de la formation du cristal ; c'est le cas quand il s'agit d'atomes très robustes comme les gaz rares avec leurs couches électroniques complètes.

L'interaction entre deux objets neutres, robustes et bien localisés, se compose de deux contributions dont l'origine physique est évidente. D'une part une forte répulsion à courte distance, chaque atome ayant par nature un domaine propre réservé résistant à toute invasion extérieure. D'autre part, l'interaction attractive de van der Waals, donnant une énergie potentielle décroissant en module comme R^{-6}. Au total, l'interaction entre deux atomes neutres a l'allure représentée sur la figure 29.1.

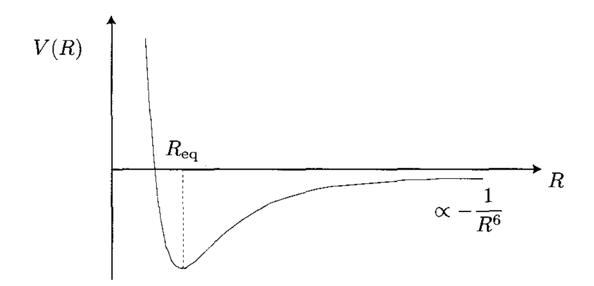


Figure 29.1: Allure de l'énergie potentielle d'interaction entre deux objets neutres.

Autant la forme à longue distance est parfaitement déterminée, et en un sens universelle, autant la répulsion de cœur varie d'un élément à l'autre. La modélisation précise de cette dernière n'a pas beaucoup d'importance (ce qui importe, c'est d'inclure de fait une répulsion), puisqu'elle intervient sur une petite échelle de distance, très peu

explorée pour le problème ici analysé[11]. Le choix d'une dépendance en R ou d'une autre est largement dicté par le souci de la simplicité analytique ; on prend couramment[12] le *potentiel 6-12* de Lennard - Jones :

$$\boxed{V(R) = 4v_0 \left[\left(\frac{d}{R}\right)^{12} - \left(\frac{d}{R}\right)^{6}\right] \qquad (v_0 > 0)} \tag{29.3}$$

La distance d'équilibre pour une paire unique est $R_{\text{eq}} = 2^{1/6}\, d \simeq 1,12d$, ce qui donne le sens physique du paramètre d dans (29.3) ; quant à l'énergie de dissociation d'une paire, elle vaut $V(R_{\text{eq}})$; avec le calibrage (29.3), on a tout simplement $V(R_{\text{eq}}) = -v_0$.

Les deux paramètres phénoménologiques d et v_0 s'obtiennent le plus souvent en mesurant, en phase vapeur, l'écart[13] à la loi $PV = Nk_{\text{B}}T$ et en prenant appui sur le développement classique du Viriel, donnant l'équation d'état d'un fluide en puissances *entières* de la densité ρ de particules ($\rho = \frac{N}{V}$). Cette formalisation en série induit un développement du même type pour toutes les grandeurs thermodynamiques ; en particulier, la pression P apparaît maintenant sous la forme :

$$P = N\,\frac{k_{\text{B}}T}{V}\,\left[1 - B(T)\,\rho - C(T)\,\rho^2 - \ldots\right] \ , \tag{29.4}$$

où les fonctions $B(T)$, $C(T)$, ..., sont par définition les *coefficients du Viriel*[14]. Le coefficient B (comme les autres) dépend de la température, mais pas de la densité ρ : ceci permet de comprendre que les valeurs obtenues par l'étude du *gaz* conservent toute leur utilité pour l'étude de la phase condensée. Les valeurs numériques de la table 29.1 montrent bien la fragilité des molécules ainsi constituées, qui ne peuvent avoir qu'une existence éphémère à l'ambiante ($(k_{\text{B}}T)_{300\,\text{K}} \simeq 25\,\text{meV}$). Quoi qu'il en soit, c'est aussi parce que les composants sont très peu altérés lors de la condensation que l'on peut utiliser ces paramètres déduits de l'analyse en phase vapeur pour l'énergie de cohésion du solide correspondant.

[11] Tôt ou tard, des exponentielles de Boltzmann du genre $\text{e}^{-\beta V(R)}$ interviennent dans des intégrales ; à très courte distance et aux températures usuelles, ces exponentielles tendent vers zéro à toute vitesse, puisque $V(R)$ diverge, et viennent donc effacer toute contribution à l'intégrale ; c'est pourquoi la forme précise de $V(R)$ à courte distance est en pratique sans grande importance.

[12] Le 12 n'est finalement choisi que par commodité ($2 \times 6 = 12$!), aucune raison théorique ne le justifie. Les calculs quantiques avec les fonctions d'onde donnent plutôt une "divergence" exponentielle du genre $\text{e}^{2\alpha(R_{\text{eq}}-R)}$ ($R < R_{\text{eq}}$), qui justifie l'usage du potentiel de Morse (voir problème 28.4.6) :

$$V(x) = -V_0\left[\text{e}^{-2\alpha(R-R_{\text{eq}})} - 2\text{e}^{-\alpha(R-R_{\text{eq}})}\right] \ . \tag{29.2}$$

[13] À partir du moment où on met une interaction $V(R)$ entre deux atomes du gaz, celui-ci n'est plus parfait, de sorte que son équation d'état s'écarte nécessairement de celle des gaz parfaits.

[14] Le premier terme de (29.4) redonne la loi des gaz parfaits ; les fonctions $B(T)$, $C(T)$, ... s'expriment à l'aide d'intégrales impliquant justement l'interaction entre deux atomes ; par exemple [231] :

$$B(T) = 2\pi \int_0^{+\infty} r^2 \text{d}r\,[\text{e}^{-\beta V(r)} - 1] \ . \tag{29.5}$$

La détermination de la quantité $B(T)$ (*via* la mesure des isothermes) permet de caractériser $V(R)$ et donne donc accès à v_0 et d.

Table 29.1: Valeurs de v_0 et R_{eq} pour les gaz rares

v_0, R_{eq}	Ne	Ar	Kr	Xe
$v_0^{(\text{exp})}$ (eV)	0,003 1	0,010 4	0,014 0	0,020 0
$R_{\text{eq}}^{(\text{exp})}$ (Å)	2,74	3,40	3,65	3,98

L'interaction d'une paire étant précisée, il est possible de trouver l'énergie de cohésion d'un *cristal*, construit en disposant aux nœuds d'un réseau les atomes en interaction par $V(\vec{R})$. Si $\vec{R}_i$ repère un point de ce réseau, l'interaction totale est :

$$E = \frac{1}{2} \sum_{\vec{R}_i,\, \vec{R}_j,\, i \neq j} V(\vec{R}_i - \vec{R}_j) \ , \tag{29.6}$$

où le facteur $\frac{1}{2}$ évite de compter deux fois l'interaction au sein d'une même paire. Dans la limite d'un cristal macroscopique contenant N ($\sim 10^{23}$) atomes, il y a une quasi-invariance par translation (discrète), l'énergie E devient une grandeur extensive[15] et il vient simplement, aux termes de surface près :

$$E \to \frac{N}{2} \sum_{\vec{R} \neq \vec{0}} V(\vec{R}) \qquad (N \ggg 1) \ . \tag{29.7}$$

Dans cette limite, on peut donc définir une énergie de cohésion par particule, ε :

$$\boxed{\varepsilon \stackrel{\text{déf}}{=} \lim_{N \to +\infty} \frac{E}{N} = \frac{1}{2} \sum_{\vec{R}} V(\vec{R})} \tag{29.8}$$

Si l'on choisit le potentiel (29.3), il faut donc calculer $\varepsilon = 2v_0 \sum_{\vec{R}} \left[\left(\frac{d}{R}\right)^{12} - \left(\frac{d}{R}\right)^6\right]$, où la somme court sur tous les nœuds du réseau, chacun d'entre eux étant caractérisé par un vecteur $\vec{R}$. Par exemple, pour un réseau cubique de côté a, on doit calculer des sommes, dites *sommes de réseau*, du genre :

$$\left(\frac{d}{a}\right)^n \sum_{m_1, m_2, m_3} \frac{1}{\|m_1 \vec{i} + m_2 \vec{j} + m_3 \vec{k}\|^n} \stackrel{\text{déf}}{=} \left(\frac{d}{a}\right)^n A_n \tag{29.9}$$

avec $n = 6$, 12 et où les m_i sont des entiers naturels, A_n étant un nombre pur ; malgré la valeur élevée de l'exposant, le calcul de telles sommes (triples !) n'est pas simple et des méthodes astucieuses ont été mises au point pour accélérer la convergence numérique.

[15]L'argument ne vaut que dans la mesure où les forces ne sont pas à portée infinie : de ce fait, les effets de surface/volume s'évanouissent à la limite thermodynamique. En effet, si ξ désigne la portée des interactions, le nombre d'atomes de surface en interaction mutuelle est $\sim (N/L^3)\,\xi\, L^2$, pour un échantillon cubique de côté L, soit $N\,(\xi/L)$. Dans la limite thermodynamique, pour une grandeur extensive comme l'énergie, les corrections de surface s'annulent comme L^{-1}, soit comme $N^{-1/3}$ lorsque la densité est maintenue constante.

Dans le cas d'un réseau quelconque, il est commode de poser $\vec{R} = l\vec{\rho}(\vec{R})$, où l est la distance entre plus proches voisins. Pour un réseau cubique à faces centrées (CFC), le cas des gaz rares, les sommes de réseau donnent :

$$\boxed{A_6 \simeq 14,45 \qquad A_{12} \simeq 12,13 \qquad \text{(CFC)}} \tag{29.10}$$

En l'état actuel des choses, l'énergie du réseau par particule est :

$$\varepsilon = 2v_0 \left[A_{12} \left(\frac{d}{l}\right)^{12} - A_6 \left(\frac{d}{l}\right)^{6}\right] . \tag{29.11}$$

Par minimisation, on trouve la distance d'équilibre l_{eq}, reliée au côté de la maille cubique conventionnelle du cristal CFC constitué ($a_{eq} = \sqrt{2}\, l_{eq}$) :

$$l_{eq} = \left(\frac{2A_{12}}{A_6}\right)^{1/6} d \simeq 1,09\,d \ , \qquad a_{eq} \simeq 1,54\,d \ . \tag{29.12}$$

La distance l_{eq} est un peu inférieure à la valeur d'équilibre d'une seule paire ($\simeq 1,12\,d$, voir juste après (29.3)), traduisant l'augmentation de compacité quand on passe en phase solide : une grande quantité d'atomes qui tous s'attirent ont évidemment tendance à former un assemblage relativement plus ramassé. Les valeurs numériques sont en assez bon accord avec l'expérience, comme le montre la table 29.2. On constate d'ailleurs que l'accord est d'autant meilleur que l'atome est plus lourd : l'oubli des termes de point-zéro est bien de moins en moins préjudiciable quand la masse des atomes augmente. D'une façon quasi-générale, les vibrations irréductibles d'après Heisenberg jouent le rôle d'une répulsion effective : quand les atomes vibrent, ils occupent plus de place et leur distance d'équilibre se trouve, de fait, augmentée.

Table 29.2: Valeurs de l_e et R_{eq} pour les gaz rares

v_0, R_{eq}	Ne	Ar	Kr	Xe
$l_{eq}^{(exp)}$ (Å)	3,13	3,75	3,99	4,33
$l_{eq}^{(th)} = 1,09\,d$ (Å)	2,99	3,71	3,98	4,34
$\varepsilon_{eq}^{(exp)}$ (eV)	−0,02	−0,08	−0,11	−0,17
$\varepsilon_{eq}^{(th)} = -8,60\,v_0$ (eV)	−0,027	−0,089	−0,120	−0,172

Une fois obtenue la distance d'équilibre, on en déduit l'énergie de cohésion (toujours comptée par atome) en reportant l'expression (29.12) de l_{eq} dans (29.11) :

$$\boxed{\varepsilon_{eq} \overset{\text{déf}}{=} \varepsilon(l_{eq}) = -v_0\,\frac{A_6^2}{2A_{12}}} \tag{29.13}$$

qui donne, pour le CFC, $\varepsilon_{eq} \simeq -8,60\,v_0$. L'accord est convenable, comme le montre la table 29.2 ; à nouveau, il est d'autant meilleur que l'atome est plus massif. Les énergies

expérimentales sont plus élevées (en valeurs algébriques, l'énergie de liaison est plus faible, la liaison est donc plus fragile) puisque l'on a négligé une contribution *positive* à l'énergie, à savoir l'énergie de vibration[16].

La faiblesse de la liaison, déjà relevée pour la phase vapeur, contraste avec l'augmentation importante de la liaison par atome à la suite de la condensation : il y a presque un facteur 10 entre ε_{eq} du solide et v_0, énergie de liaison d'une molécule en phase vapeur. Ainsi, sans rien changer aux interactions microscopiques, la coopération de tous les atomes au sein du solide donne à celui-ci une énergie de cohésion tout à fait respectable, bien qu'elle reste encore très inférieure à celle d'un cristal ionique ou métallique, qui est toujours de l'ordre de plusieurs eV (voir plus loin).

29.2.3 Cohésion des réseaux ioniques

Pour un cristal ionique, l'interaction dominante à une distance $\gtrsim a_0$ est celle de Coulomb, qui décroît très lentement comme R^{-1}. Il existe toujours une répulsion à courte distance $V_{\text{cœur}}$, de sorte qu'il faut écrire (dans la limite macroscopique $N \ggg 1$) :

$$E = \frac{N}{2} \sum_{\vec{R}} [V_{\text{cœur}}(\vec{R}) + V_{\text{Coulomb}}(\vec{R})] \equiv E_{\text{cœurs}} + E_{\text{Coulomb}} \quad . \tag{29.14}$$

En raison de la décroissance lente de l'interaction de Coulomb nue, le calcul de la somme impliquant V_{Coulomb} est toujours délicat. La difficulté est visible immédiatement pour un réseau à une dimension, où alternent des ions positifs et des ions négatifs, portant chacun une charge élémentaire et situés à la distance a les uns des autres. L'énergie potentielle de l'ion situé à l'origine compte tenu de tous les autres est :

$$2\frac{e'^2}{a}\left(-1 + \frac{1}{2} - \frac{1}{3} + \ldots \right) , \tag{29.15}$$

le facteur 2 global prenant en compte les deux moitiés du réseau, à gauche et à droite de l'origine. La série harmonique alternée, $\sum_{n\in\mathbb{N}}(-1)^n/n$, est convergente mais approche sa somme $-\ln 2$ avec une lenteur désespérante. D'ailleurs, on pourrait décider de compter autrement, en prenant d'abord l'interaction entre l'ion à l'origine avec tous ses partenaires de même signe ; ainsi viendrait la première série $2(e'^2/a)\sum_{n\in\mathbb{N}} 1/(2n)$. L'interaction entre le même ion central et ceux de charges opposée est $-2(e'^2/a)\sum_{n\in\mathbb{N}} 1/(2n+1)$. Le seul ennui est que ces deux séries sont … divergentes ! La difficulté s'aggrave encore dans le cas tridimensionnel, mais des méthodes très ingénieuses ont été inventées pour calculer ces séries non absolument convergentes[17].

[16] L'énergie de vibration contient deux termes (cinétique et potentiel), tous deux positifs.

[17] À la réflexion, l'argumentation peut paraître un peu spécieuse, car la Nature se moque de savoir comment le Physicien s'y prend pour faire ses calculs. Et, en définitive, comment peut-on être sûr que le calcul, fait d'une façon ou d'une autre, donne le bon résultat ? En réalité, il faut bien comprendre ceci : quoi qu'il en soit, les cristaux sont toujours de taille *finie*, contenant N ions. Alors, quel que soit N, aussi gigantesque soit-il, le résultat est unique, indépendant de la procédure de sommation utilisée. Finalement, s'agissant de calculer la somme d'un grand nombre de termes, c'est le souci d'une

Dimensionnellement parlant, l'énergie de Coulomb est forcément de la forme :

$$E_{\text{Coulomb}} \propto -N\frac{e'^2}{l} \ , \tag{29.16}$$

l étant à nouveau la distance entre deux plus proches voisins. En comptant par paire d'ions (l'habitude pour les cristaux ioniques), on définit précisément une constante A, dite *constante de Madelung*, par la relation :

$$E_{\text{Coulomb}} \stackrel{\text{déf}}{=} -A\frac{N}{2}\frac{e'^2}{l} \ ; \tag{29.17}$$

A est un nombre, qui est l'enjeu des différentes méthodes de sommation. Par paire d'ions, (29.17) donne :

$$\boxed{\varepsilon_{\text{Coulomb}} = -A\frac{e'^2}{a}} \tag{29.18}$$

La constante de Madelung est un nombre voisin de 1, comme le montre la table 29.3.

Table 29.3: Valeurs de la constante de Madelung pour quelques structures cristallines

Structure cristalline	Constante de Madelung
CsCl	1,7627
NaCl	1,7476
Pechblende	1,6381

L'expression (29.18) est une borne supérieure de l'énergie de liaison, puisque la répulsion de cœur (déstabilisante) n'est pas encore prise en compte. En adoptant comme précédemment une loi-puissance pour cette dernière, l'énergie de cohésion dans le modèle à ions fixes est :

$$\varepsilon = -A\frac{e'^2}{a} + \frac{C}{a^m} \qquad (N \gg 1) \ ; \tag{29.19}$$

par minimisation on trouve la distance d'équilibre $a_{\text{eq}} = \left[mC/(Ae'^2)\right]^{1/(m-1)}$, d'où l'on déduit l'énergie par paire d'ions $\varepsilon_{\text{eq}} \stackrel{\text{déf}}{=} \varepsilon(a_{\text{eq}})$:

$$\boxed{\varepsilon_{\text{eq}} = -A\frac{m-1}{m}\frac{e'^2}{a_{\text{eq}}}} \tag{29.20}$$

Pour m assez grand, ce dernier résultat est très peu différent de $-A\,(e'^2/a_{\text{eq}})$: a_{eq} dépend faiblement de m. Pour les cristaux ioniques, le cœur fortement répulsif joue peu de rôle sur le plan purement énergétique : en raison de son intensité et de sa longue portée, l'interaction de Coulomb est de loin l'interaction dominante. Comme a_{eq} est de l'ordre de quelques angströms, ε_{eq} vaut typiquement plusieurs eV.

bonne efficacité numérique qui dicte le choix d'une méthode ou d'une autre. Physiquement, la réponse est satisfaisante si, une fois la somme calculée par la méthode la plus rapide, l'énergie devient – à l'approximation souhaitée – une grandeur extensive.

29.2.4 Cohésion des métaux

En première approximation, un métal peut se représenter comme un réseau d'ions tous positifs noyés dans une mer d'électrons presque libres, mais néanmoins soumis au potentiel périodique du réseau. Ce sont ces électrons, libérés lors de la formation du cristal (dont les fragments sont neutres au départ), qui constituent la colle assurant la cohésion du cristal – autrement, tous les ions positifs se repoussant, l'édifice ne pourrait être stable et exploserait. Dans la théorie des métaux, il est d'usage d'introduire le rayon de la sphère moyenne occupée par un électron, noté r_s, défini comme[18] :

$$\boxed{V \stackrel{\text{déf}}{=} N\frac{4\pi r_s^3}{3}} \tag{29.22}$$

où N est le nombre total d'électrons et V le volume de l'échantillon ; pour les métaux ordinaires, r_s vaut quelques a_0.

L'énergie de cohésion se définit par référence à l'énergie des constituants séparés. Dans le cas des métaux, des électrons sont libérés lors de la formation du solide et se retrouvent presque libres (dans le cas d'un cristal ionique, chaque électron libéré par l'atome d'une espèce est piégé par un atome de l'autre). L'énergie du cristal constitué doit donc incorporer l'énergie d'interaction entre les ions (comme avant), mais également toutes les énergies relatives aux électrons (énergie potentielle et énergie cinétique).

On admet que l'énergie potentielle est celle résultant de l'interaction de Coulomb nue[19] ; elle donne lieu à plusieurs sortes de contributions : interactions entre ions, répulsives, interactions entre électrons, également répulsives, interactions entre les ions et les électrons, attractives. Délaissant toujours les vibrations des ions positifs, ceux-ci se bornent à jouer le rôle de charges ponctuelles fixes immergées dans un fond continu négatif compensateur (les électrons), le système étant globalement neutre. Naïvement, le calcul peut se faire suivant les mêmes techniques que pour les réseaux ioniques, et on trouve, par atome, dans le cas d'un réseau cubique centré :

$$\varepsilon_{\text{Coulomb}} = -\frac{24,35}{r_s/a_0}\,\text{eV} \ . \tag{29.23}$$

Le résultat (29.23) est purement classique et incorpore toutes les interactions coulombiennes *classiques* de la distribution de charge (électrons et ions) ainsi définie (réseau d'ions fixes, de structure déterminée et distribution électronique uniforme et continue).

En fait, le calcul du terme de répulsion entre électrons ne se réduit pas au terme classique inclus dans (29.23) qui, par nature, ignore le Principe de Pauli. Dans

[18] La définition de r_s fluctue d'un auteur à l'autre ; l'autre définition usuelle donne un r_s adimensionné (on compte le rayon de la sphère en unités a_0) :

$$V \stackrel{\text{déf}}{=} N\frac{4\pi r_s^3}{3}\,a_0^3 \ . \tag{29.21}$$

[19] Cette hypothèse est très critiquable, car la haute densité d'électrons conduit forcément à des phénomènes d'écrantage.

l'approximation minimale prenant ce dernier en compte pour des fermions (une fonction d'onde en forme de déterminant de Slater), apparaissent en plus des termes d'échange (comme pour l'atome d'hélium) donnant au total une contribution toujours négative pour deux électrons de même spin (trou de Fermi). Ceci réduit fortement la répulsion de Coulomb (qui est déstabilisante) par rapport à la situation *spinless*[20] ; au total, la prise en compte de ces termes abaisse la moyenne de l'énergie potentielle totale de façon importante. Le calcul de la contribution d'échange par perturbation donne très précisément[21] :

$$\varepsilon_{\text{échange}} = -\frac{3}{2\pi}\left(\frac{9\pi}{4}\right)^{1/3}\frac{a_0}{r_s}\frac{e'^2}{2a_0} \simeq -\frac{12,47}{r_s/a_0}\,\text{eV} \ , \tag{29.24}$$

de sorte que la prise en compte (rustique) de l'indiscernabilité des électrons en interaction donne au total l'énergie de Coulomb :

$$\varepsilon_{\text{Coulomb}} \simeq -\frac{36,82}{r_s/a_0}\,\text{eV} \ . \tag{29.25}$$

Cela étant, à cause de la faible inertie des électrons et de leur mouvement presque libre au sein du métal, il faut introduire leur énergie cinétique, qui entre bien dans le bilan d'énergie par rapport aux constituants (atomes neutres) séparés, où l'énergie cinétique des électrons, alors *liés*, est prise en compte dans l'énergie de constitution des atomes à l'état libre (en phase gazeuse). En raison de la rusticité du modèle en cours on calcule cette énergie cinétique comme si les électrons étaient complètement libres – ce qui suppose que l'énergie potentielle de chaque électron dans le champ des ions positifs est une petite correction[22]. L'énergie de l'un d'entre eux est alors :

$$\varepsilon(\vec{k}) = \frac{\hbar^2\vec{k}^2}{2m} \ , \tag{29.26}$$

où $\vec{k}$ est le vecteur d'onde. Pour un alcalin, il y a un électron presque libre par atome, c'est-à-dire que s'il y a N ions dans le cristal il y a aussi N électrons, qu'il faut répartir dans les états à une particule conformément au Principe de Pauli. Les électrons étant considérés comme libres, et adoptant, tout comme pour l'effet photoélectrique (chapitre 24, sous-section 24.5.2) des conditions cycliques de Born - von Kármán dans une boîte cubique de côté L, ils sont représentés par des ondes planes $\propto \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$; les valeurs de $\vec{k}$ sont quantifiées, chaque composante étant un multiple entier de $\frac{2\pi}{L}$, l'élément de volume dans l'espace réciproque étant un petit cube infinitésimal de côté $\frac{2\pi}{L}$.

[20]En effet, pour la moitié des paires d'électrons, la répulsion est fortement diminuée par le trou de Fermi.

[21]Ce calcul n'est pas difficile, mais n'est pas pour autant élémentaire. Il passe par le calcul de l'énergie à deux corps dans une approximation de type déterminant de Slater ou, ce qui ici revient au même comme on peut s'en convaincre, un calcul de perturbation au premier ordre. Le problème 29.4.2 est consacré à cette question.

[22]Cette idée est le fondement de l'approximation dite des *électrons presque libres*, qui sera exposée ultérieurement (voir chapitre 30, sous-section 30.4.1). Notons toutefois que même si l'effet du réseau est quantitativement petit, il provoque néanmoins une modification qualitative spectaculaire de la loi de dispersion de l'électron : la parabole libre devient une succession d'arcs déformés et disjoints, d'où l'apparition de *gaps* d'énergie (voir problème 17.8.6, et figure 30.5).

Le principe de Pauli exige de répartir les électrons dans ces états mettant au plus 2 électrons par état orbital, compte tenu du spin. En procédant ainsi, on remplit ceux-ci jusqu'à un certain vecteur d'onde k_{F}, appelé vecteur d'onde de Fermi. Dans la sphère de rayon k_{F}, il y a $\frac{4\pi}{3}\,k_{\mathrm{F}}^3/\left(\frac{2\pi}{L}\right)^3$ petits cubes et chacun peut accepter deux électrons ; pour un alcalin, il vient ainsi $2\frac{4\pi}{3}\,k_{\mathrm{F}}^3/\left(\frac{2\pi}{L}\right)^3 = N$, soit :

$$\boxed{k_{\mathrm{F}} = \frac{2\pi}{L}\left(\frac{3N}{8\pi}\right)^{1/3} \propto \left(\frac{N}{V}\right)^{1/3}} \tag{29.27}$$

Cela étant précisé, il est possible de calculer l'énergie cinétique de ces N électrons, dont l'expression est :

$$E_{\mathrm{cin}} = \sum_{\mathrm{spin}}\sum_{k\,\leq\,k_{\mathrm{F}}} \frac{\hbar^2\vec{k}^{\,2}}{2m} \quad\longrightarrow\quad 2\left(\frac{L}{2\pi}\right)^3 \int_0^{k_{\mathrm{F}}} 4\pi k^2 \mathrm{d}k\,\frac{\hbar^2 k^2}{2m} \ ; \tag{29.28}$$

l'intégrale est élémentaire et donne :

$$\boxed{E_{\mathrm{cin}} = N\,\frac{3}{5}\,\frac{\hbar^2 k_{\mathrm{F}}^2}{2m} \overset{\mathrm{déf}}{=} N\,\frac{3}{5}\,\varepsilon_{\mathrm{F}}} \tag{29.29}$$

la dernière égalité définissant l'énergie de Fermi ε_{F} pour le gaz d'électrons libres. En introduisant la longueur r_s définie en (29.22), on trouve finalement l'énergie cinétique par électron sous la forme :

$$\varepsilon_{\mathrm{cin}} = \frac{3}{5}\,\frac{e'^2}{2a_0}\left(\frac{9\pi}{4}\right)^{2/3}\left(\frac{a_0}{r_s}\right)^2 \quad\Longleftrightarrow\quad \varepsilon_{\mathrm{cin}} \simeq \frac{30,07}{(r_s/a_0)^2}\,\mathrm{eV}\ . \tag{29.30}$$

Notons que comme k_{F} augmente avec la densité N/V, il en va de même de E_{cin} : plus le gaz d'électrons est dense, plus grande est l'énergie cinétique.

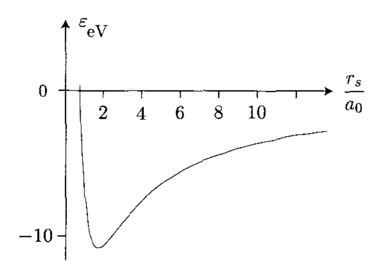


Figure 29.2: Variation de l'énergie par atome (par électron) pour un métal alcalin (voir (29.31)).

En additionnant maintenant toutes les contributions à l'énergie (29.23), (29.30) et (29.24), on obtient l'énergie de cohésion par atome, pour un alcalin :

$$\boxed{\varepsilon_{\mathrm{eV}} = \frac{30,07}{(r_s/a_0)^2} - \frac{36,82}{r_s/a_0} \equiv T(r_s) + V(r_s)} \tag{29.31}$$

La minimisation par rapport à r_s donne $r_{s\,\mathrm{eq}} \simeq 1,6\,a_0$, d'où l'énergie de cohésion :

$$\boxed{|\varepsilon(r_{s\,\mathrm{eq}})| \simeq 11,3\,\mathrm{eV}} \tag{29.32}$$

Ces résultats sont assez médiocres, les valeurs expérimentales de $\frac{r_s}{a_0}$ variant en fait entre 2 et 6. Ils donnent néanmoins l'occasion, au-delà du désaccord numérique[23], de commenter physiquement le résultat en partant de l'idée qu'il reste qualitativement correct pour l'essentiel.

La dépendance en r_s se comprend bien : le terme cinétique provient d'opérateurs donnant un facteur d'échelle λ^{-2} vis-à-vis de tout *scaling* des longueurs, d'où le r_s^{-2} ; le même argument pour l'énergie potentielle (purement de Coulomb) donne une dépendance en r_s^{-1}. Comme on l'a vu au chapitre 27 (section 27.5), la variation de r_s est équivalente à un *scaling*, la minimisation conduisant automatiquement à la satisfaction du théorème du Viriel ; numériquement, on a bien :

$$\boxed{2T(r_{s\,\mathrm{eq}}) + V(r_{s\,\mathrm{eq}}) = 0} \tag{29.33}$$

Clairement, c'est du côté des hautes densités électroniques que le calcul effectué ci-dessus est *a priori* valide ; en effet, lorsque r_s est petit, c'est l'énergie cinétique $\varepsilon_{\mathrm{cin}}$ qui est dominante et la prise en compte élémentaire (i.e. par perturbation) de l'interaction coulombienne concerne alors des termes énergétiques relativement petits. À l'autre extrême (faibles densités, grand r_s), le calcul précédent tend à montrer (malgré ses limites) que l'énergie potentielle est dominante, l'énergie cinétique étant petite (les électrons sont peu mobiles). En 1934, par un tout autre moyen, Wigner [232] a prévu l'existence d'un cristal d'électrons (*cristal de Wigner*) ; cette prévision théorique a reçu une confirmation expérimentale en 1979 [233], par la mesure de la modification des ondes capillaires de l'hélium liquide, à la surface duquel un gaz d'électrons avait été déposé. Les altérations observées pour la loi de dispersion de ces ondes sont compatibles avec l'existence d'un réseau bidimensionnel triangulaire d'électrons.

29.3 Ordre spatial d'un cristal. Diffraction

Dans cette section sont présentées quelques notions indispensables pour la bonne compréhension des traitements formels entrepris par la suite. Il s'agit en fait de donner quelques rudiments concernant la description précise de la structure physique des solides cristallisés, caractérisés par un ordre à portée infinie, et d'examiner les premières conséquences physiques de cet arrangement supposé parfait.

[23] D'une façon générale, la théorie de perturbation ordinaire pour traiter l'interaction de Coulomb entre électrons n'est pas satisfaisante : le calcul à des "ordres supérieurs" fait apparaître des termes singuliers du genre $r_s \ln r_s$.

29.3.1 Réseaux. Structure physique d'un réseau

Soit, dans $\mathbb{R}^D$, D vecteurs linéairement indépendants $\vec{a}_i$, $i = 1, 2, ..., D$. On appelle réseau de Bravais[24], noté $\mathcal{B}$, l'ensemble des points (*nœuds*) dont le rayon-vecteur $\vec{R}$ est de la forme :

$$\vec{R} = \sum_{i=1}^{D} n_i \vec{a}_i \qquad (n_i \in \mathbb{Z}) \ ; \tag{29.34}$$

le point important dans cette définition est que $\vec{R}$ est une combinaison linéaire en *entiers* des vecteurs $\vec{a}_i$ – c'est pourquoi un réseau de Bravais en dimension D est parfois noté $\mathbb{Z}^D$. Les vecteurs $\vec{a}_i$ sont appelés *vecteurs primitifs*, puisqu'ils permettent d'engendrer le réseau par la relation de définition (29.3.1) ; un réseau étant donné, leur choix n'est visiblement pas unique. À l'ensemble des vecteurs $\vec{R}$, est associé un ensemble de translations $T(\vec{R})$, appelé translations du réseau (toute translation du réseau relie deux points du réseau) que l'on peut visiblement munir d'une structure de groupe par la loi de composition interne $T(\vec{R})T(\vec{R}') = T(\vec{R} + \vec{R}')$. L'analyse systématique montre qu'il existe en tout et pour tout 14 réseaux de Bravais distincts (cubique, quadratique, orthorhombique, monoclinique, triclinique), chacun pouvant exister sous différentes formes (par exemple, pour le cubique : simple (C), centré (CC) et à faces centrées (CFC)).

Évidemment, outre la symétrie de translation, un réseau peut posséder d'autres éléments de symétrie (rotation, réflexion dans un plan, inversion d'espace, etc.), qui se distinguent des translations en ce sens qu'elles possèdent des points fixes (invariants). L'ensemble de ces opérations constitue la symétrie dite *ponctuelle*[25], afin de la distinguer de la symétrie de translation.

La *maille primitive* d'un réseau de Bravais est le volume de $\mathbb{R}^D$ qui, soumis à *toutes* les translations du réseau remplit complètement l'espace sans aucun recouvrement : c'est la brique élémentaire (incompressible !) *pavant* l'espace sans laisser le moindre trou ; une maille primitive (le choix n'est pas unique) contient exactement un point du réseau et un seul[26], et toutes ont le même volume v. Pour un réseau tridimensionnel, v est égal au module du produit mixte construit sur trois vecteurs primitifs :

$$\boxed{v = |(\vec{a}_1, \vec{a}_2, \vec{a}_3)|} \tag{29.35}$$

[24] Le mot réseau désigne plus généralement un ensemble infini de points régulièrement répartis dans l'espace suivant une règle précise. Tous les réseaux ne sont pas des réseaux de Bravais, l'exemple le plus connu étant le réseau en nid d'abeille.

[25] le qualificatif rappelle qu'il existe des points fixes.

Un point mérite d'être mentionné : la symétrie d'ordre 5 (une rotation de $\frac{2\pi}{5}$ par exemple) est incompatible avec la symétrie de translation d'un réseau de Bravais. On a toutefois découvert assez récemment [234] des matériaux remarquables en ce sens qu'ils exhibent une symétrie pentagonale, tout en donnant des clichés de diffraction à l'instar d'un réseau de Bravais. Ces systèmes, appelés *quasi-cristaux*, semblent donc à première vue invalider le dogme de la cristallographie interdisant la symétrie d'ordre 5.

[26] Il est d'usage courant, pour la commodité des représentations géométriques, de manipuler des mailles qui ne sont pas primitives au sens où on peut reproduire tout le réseau en leur appliquant seulement un *sous*-ensemble des translations. De telles mailles sont appelées mailles *conventionnelles*. Par exemple, pour le cubique centré, une telle maille est constituée par le cube de côté a, munie de son atome central ; cette maille conventionnelle est deux fois plus grande en volume que toute maille primitive du CC puisqu'elle contient deux points du réseau.

Il est toujours possible de choisir comme maille primitive une maille ayant par elle-même la symétrie (ponctuelle) du réseau. Un exemple est la maille dite[27] de Wigner - Seitz, constituée des points de $\mathbb{R}^D$ qui sont plus proches d'un nœud donné que de tous les autres nœuds. De façon évidente, il s'agit d'un polyèdre dont les faces sont dans les plans médiateurs des segments joignant deux voisins du réseau de Bravais. Elle présente l'avantage d'être indépendante du choix des vecteurs primitifs (elle est construite directement sur les points du réseau) et c'est d'ailleurs pour cette raison qu'elle en possède forcément la symétrie ponctuelle.

Un réseau de Bravais est un ensemble de points. Cette définition *géométrique* acquiert un contenu physique en précisant les objets situés en ses nœuds (*décoration*). On appelle *motif* (ou *base*) l'unité physique (atome, molécule, ion, etc.) attachée à un nœud et dont la duplication à l'infini par translation reproduit le cristal physique, supposé parfait d'un point de vue structurel et chimique. Un exemple intéressant est le réseau en nid d'abeille – qui n'est pas un réseau de Bravais –, mais qui peut être construit à partir d'un réseau triangulaire muni d'une base à deux points. L'ensemble réseau et base définit complètement la *structure cristalline* d'un matériau donné.

Le réseau *réciproque* d'un réseau de Bravais apparaît spontanément dès que l'on fait de l'analyse de Fourier. En effet, l'objet fondamental de l'analyse de Fourier est l'onde plane $\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$ où $\vec{k}$ est un vecteur d'onde ; par ailleurs, on connaît l'importance du fait que $\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} = 1$: c'est la condition pour avoir des interférences constructives. On en vient ainsi tout naturellement, un réseau de Bravais $\mathcal{B} \equiv \{\vec{R}\}$ étant donné, à considérer l'ensemble des points associés à des vecteurs d'onde, notés $\vec{K}$ pour les distinguer, satisfaisant la relation $\mathrm{e}^{\mathrm{i}\vec{K}.\vec{R}} = 1$; cet ensemble de points est plongé dans $\mathbb{R}^D$, mais ce n'est pas l'espace réel – les vecteurs correspondants $\vec{K}$ sont homogènes à l'inverse d'une longueur, comme un vecteur d'onde : c'est l'*espace réciproque*. Il est clair que l'ensemble des extrémités de ces vecteurs $\vec{K}$ ne forme pas un *continuum* mais a la nature discrète d'un réseau : c'est le *réseau réciproque* de $\mathcal{B}$, noté $\tilde{\mathcal{B}}$, qui est donc l'ensemble des points satisfaisant par définition[28] :

$$\boxed{\tilde{\mathcal{B}} \stackrel{\text{déf}}{=} \{\vec{K}\} \ , \ \mathrm{e}^{\mathrm{i}\vec{K}.\vec{R}} = 1 \qquad \forall \vec{R} \in \mathcal{B}} \tag{29.36}$$

De la relation de définition (29.36), il résulte de suite : $\forall \vec{r} \in \mathbb{R}^D$, $\mathrm{e}^{\mathrm{i}\vec{K}.(\vec{R}+\vec{r})} = \mathrm{e}^{\mathrm{i}\vec{K}.\vec{r}}$.

Une autre définition (équivalente) du réseau réciproque passe par un choix de vecteurs primitifs $\{\vec{a}_i\}$ du réseau direct $\mathcal{B}$, supposé à trois dimensions pour fixer les idées ; soit alors les trois vecteurs $\vec{b}_i$ définis comme suit :

$$\vec{b}_1 \stackrel{\text{déf}}{=} 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1.(\vec{a}_2 \times \vec{a}_3)} \ , \quad \vec{b}_2 \stackrel{\text{déf}}{=} 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_2.(\vec{a}_3 \times \vec{a}_1)} \ , \quad \vec{b}_3 \stackrel{\text{déf}}{=} 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_3.(\vec{a}_1 \times \vec{a}_2)} \ , \tag{29.37}$$

tous les dénominateurs étant en fait égaux entre eux, en raison de l'invariance du produit mixte par permutation circulaire. Ces égalités étant posées, on voit immédiatement que

[27] Dans la littérature mathématique des réseaux de points, la maille de Wigner - Seitz s'appelle maille (ou cellule) de Voronoï. Cette maille a été introduite en premier par Brillouin pour le réseau réciproque.

[28] Une fois introduit le réseau réciproque d'un réseau de Bravais, ce dernier est parfois appelé réseau *direct*, pour la symétrie du langage.

tout vecteur $\vec{K}$ combinaison en entiers des $\vec{b}_i$ satisfait l'égalité de définition (29.36). On note aussi que $\vec{a}_1.\vec{b}_1 = 2\pi$ et que $\vec{a}_1.\vec{b}_2 = 0, \ldots,$ soit plus généralement :

$$\boxed{\vec{a}_i.\vec{b}_j = 2\pi\,\delta_{ij}} \tag{29.38}$$

Tout comme le produit mixte $\vec{a}_1.(\vec{a}_2 \times \vec{a}_3)$ est, au signe près, le volume d'une maille primitive de $\mathcal{B}$, le volume $\tilde{v}$ d'une maille primitive de $\tilde{\mathcal{B}}$ est égal à $|(\vec{b}_1, \vec{b}_2, \vec{b}_3)|$, et on a[29] :

$$\boxed{v\,\tilde{v} = (2\pi)^3} \tag{29.39}$$

le produit des volumes étant bien un nombre pur ($[v]$ = [longueur^3], $[\tilde{v}]$ = [longueur]$^{-3}$). Enfin, il est facile de voir que le réseau réciproque est lui-même un réseau de Bravais, un fait anticipé par la notation $\tilde{\mathcal{B}}$, et que le réseau réciproque de $\tilde{\mathcal{B}}$ est le réseau (direct) $\mathcal{B}$: $\tilde{\tilde{\mathcal{B}}} = \mathcal{B}$.

♦ *Remarque*

Il est d'usage de dire que le réseau réciproque d'un réseau de Bravais bidimensionnel est de dimension $D = 2$; toutefois, si l'on s'en tient à la définition première (29.36), il est plus logique de dire que, en pareil cas, $\tilde{\mathcal{B}}$ est constitué de droites infinies perpendiculaires au réseau direct et dont les intersections avec $\mathbb{R}^2$ constituent le réseau réciproque bidimensionnel considéré habituellement. En effet, soit un vecteur $\vec{K}$ ainsi défini dans le plan ($\vec{K} \in \tilde{\mathcal{B}}_{D=2} \equiv \mathbb{Z}^2$) et satisfaisant donc :

$$\mathrm{e}^{\mathrm{i}\vec{K}.\vec{R}} = 1 \qquad \forall \vec{R} \in \mathcal{B}_{D=2} \equiv \mathbb{Z}^2 \;; \tag{29.40}$$

si on ajoute à $\vec{K}$ (dans le plan) n'importe quel vecteur perpendiculaire $\vec{K}_\perp$ à ce plan, $\vec{K}\,' = \vec{K} + \vec{K}_\perp$, l'égalité (29.40) sera encore satisfaite. Le même argument permet de dire que le réseau réciproque d'un réseau de Bravais unidimensionnel $\mathcal{B}_{D=1} \equiv \mathbb{Z}$ de pas a est un ensemble de plans parallèles perpendiculaires à ce réseau. Le réseau direct coupe cette famille de plans en des points équidistants d'abscisses $m\,(2\pi)/a$, $m \in \mathbb{Z}$.

Ces distinctions peuvent jouer un rôle important quand la loi de conservation de l'impulsion est pertinente, comme on l'a vu au chapitre 25, section 25.4, Remarque p. 1179. ♦

29.3.2 Diffraction par un réseau

En raison de l'ordre géométrique à longue distance présent dans un cristal, il existe un très grand nombre N de diffuseurs qui vont pouvoir *coopérer* et produire des figures de diffusion extrêmement sélectives en direction, la longueur d'onde du rayonnement étant

[29] (29.39) se vérifie facilement en utilisant $(\vec{U} \times \vec{V}) \times \vec{W} = (\vec{W}.\vec{U})\,\vec{V} - (\vec{V}.\vec{W})\,\vec{U}$.

fixée. On parle alors traditionnellement de *diffraction* pour désigner cette sélectivité étant entendu que, au niveau de chaque diffuseur, il s'agit toujours de la simple diffusion élastique de la lumière au sens rappelé ci-dessous. C'est l'énormité de N allié à l'existence d'un ordre à portée infinie qui change qualitativement les phénomènes observés quand on passe d'un milieu désordonné (un gaz), ou présentant un ordre à courte distance (amorphe, liquide), à un cristal. En parlant schématiquement, la réponse du système est *floue* ou à peine structurée dans les premiers cas, et est une réponse par tout ou rien dans le dernier.

Diffusion des rayons X et des neutrons

La diffusion des rayons X et des neutrons est une méthode d'une grande efficacité pour étudier la structure de la matière, que celle-ci soit en phase diluée ou condensée. Le principe en est simple : on envoie sur le système à étudier une sonde externe parfaitement caractérisée (énergie ou longueur d'onde, polarisation[30]). L'interaction sonde - système est complexe en général ; pour simplifier, seules les interactions *élastiques* seront considérées dans la suite[31]. Dans ce cadre, la *lumière* (au sens large) issue du milieu a la même fréquence (énergie) que l'onde incidente ; son vecteur d'onde $\vec{k}_f$ a donc le même module que celui du faisceau incident, $\vec{k}_i$: $\|\vec{k}_f\| = \|\vec{k}_i\|$. En tout cas, l'interaction sonde - cible est supposée faible, puisque l'on veut une photographie de la matière telle qu'elle est, c'est-à-dire perturbée aussi peu que possible.

Rayons X Pour connaître la structure spatiale de la matière de basse énergie, dont l'échelle de longueur est de l'ordre de l'angström, il faut utiliser une lumière de longueur d'onde comparable. S'agissant de sonder la matière avec un champ électromagnétique, ce sont donc les rayons X qui conviennent ; l'énergie correspondante des photons est de quelques dizaines de keV, très supérieure aux énergies intra-atomiques ou intra-moléculaires. Le champ incident se couple à la matière essentiellement par son champ électrique, et met en vibration forcée à la même fréquence les électrons du milieu (qui lui apparaissent comme presque libres) : classiquement, c'est la diffusion Thomson (voir Tome I, problème 6.4.5), caractérisée par une section efficace $\sigma_{\mathrm{T}} = (8\pi/3)\, r_{\mathrm{e}}^2$, quasi-indépendante de la fréquence. La vibration forcée donne aux électrons une accélération non-nulle et, comme toute particule chargée accélérée rayonne, ceux-ci émettent un rayonnement électromagnétique, évidemment de même fréquence puisqu'il s'agit d'un régime forcé. Classiquement, la puissance rayonnée vers l'extérieur par une charge accélérée, $\mathcal{P}$, est donnée par la formule de Larmor (Tome I, éq. (1.19)). Quantiquement, tout ceci se

[30]Dans le cas des neutrons, la polarisation est liée au spin ($\pm\hbar/2$) des neutrons incidents ; un faisceau non polarisé est un faisceau dans lequel l'orientation du spin est aléatoire.

[31]Bien sûr, les expériences de diffusion inélastique sont aussi d'usage courant et apportent d'autres informations, notamment elles donnent accès aux lois de dispersion $\varepsilon(\vec{k})$ des excitations élémentaires du solide (phonons, magnons, etc.) ; en raison de la conservation énergie - impulsion, la mesure de la perte d'énergie $\Delta E(\vec{q})$ des projectiles à transfert de moment $\vec{q} \stackrel{\text{déf}}{=} \vec{k}_{\mathrm{f}} - \vec{k}_{\mathrm{i}}$ donné, par création d'une excitation de vecteur d'onde $\vec{k}$, permet de reconstituer point à point la loi de dispersion par l'égalité $\varepsilon(\vec{k} = -\vec{q}) = \Delta E(\vec{q})$ (voir chapitre 31, section 31.8).

transcrit en termes de section efficace de diffusion élastique des photons, le champ se couplant à la matière comme on l'a vu au chapitre 24, sous-section 24.5.1 (le formalisme semi-classique est très largement suffisant quand il s'agit de diffraction). L'effet important pour la suite – tout comme avec les neutrons – est l'addition coopérative des ondes diffusées par les atomes situés aux nœuds du réseau. Bien évidemment, et au contraire de l'interaction d'un atome avec une onde dans le visible, il ne faut pas négliger les déphasages $\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$: ici, ka_0 est précisément d'ordre 1. Ce déphasage permet de remonter à la distribution électronique d'un diffuseur (par la transformation de Fourier), et contribue de façon cruciale à la diffusion sélective.

Neutrons La diffusion des neutrons[32] relève des mêmes principes. La longueur d'onde à prendre en compte est la longueur d'onde associée de de Broglie $\lambda = \frac{h}{p} = \frac{h}{\sqrt{2ME}}$; pour les neutrons ceci s'exprime commodément comme suit :

$$\boxed{\lambda_{\text{Å}} \simeq \frac{0,286}{\sqrt{E_{\text{eV}}}}} \tag{29.41}$$

Pour un neutron thermique, $E \sim (k_{\mathrm{B}}T)_{T=300\mathrm{K}} \simeq 25$ meV, de sorte que $\lambda \sim 1,8$ Å : ce sont les neutrons thermiques qu'il convient d'utiliser pour sonder la matière ordinaire, qu'elle soit à l'état fluide ou solide. Bien sûr, le faisceau neutronique est non monochromatique et reflète l'état thermique du modérateur[33] ; si celui-ci est à la température ambiante, le maximum de la maxwellienne survient précisément pour $E \simeq 25$ meV, et il convient d'intercaler un monochromateur avant de diriger le faisceau de neutrons sur l'échantillon.

Les neutrons incidents et les constituants de la matière peuvent interagir principalement de deux façons, étant entendu que le neutron est insensible à l'interaction électromagnétique[34]. La première, relevant de l'interaction forte, est à très courte portée (la dimension nucléaire) et peut être considérée comme étant de contact[35] pour la diffusion sur des atomes ou des molécules. Par ailleurs, le neutron, grâce à son moment magnétique associé à son spin ($S = 1/2$), peut en outre se coupler aux électrons (principalement ceux qui ne sont pas appariés), auquel cas la portée de l'interaction effective est d'ordre atomique (l'échelle de longueur des fonctions d'onde de ces électrons). Enfin, toujours grâce à son moment magnétique, le neutron peut se coupler aux noyaux, et aussi aux structures magnétiques éventuellement présentes dans le matériau.

[32] Les techniques de diffusion neutronique sont nettement plus récentes que celles utilisant les rayons X ; elles ont commencé à se développer dans les années 50, lorsque des sources de neutrons suffisamment intenses sont devenues opérationnelles. Le formalisme théorique de base est principalement dû à Fermi et Schwinger.

[33] Le modérateur est là pour ralentir les neutrons : étant émis par une source à la suite de réactions nucléaires, leur énergie initiale se situe dans la gamme du MeV.

[34] en mettant de côté des effets extrêmement fins impliquant la structure de charge du neutron.

[35] Une interaction est dite *de contact* si elle est de portée nulle (les objets n'interagissent que s'ils sont au contact l'un de l'autre) ; on la représente alors par une fonction de Dirac. Une telle modélisation est bien sûr une radicalisation, qui n'a de sens que vis-à-vis d'une autre échelle préalablement définie : pour la diffusion sur la matière de basse énergie, le noyau est quasi-ponctuel – mais pas pour des collisions impliquant des neutrons de quelques MeV, utilisés pour sonder le noyau lui-même, et en savoir plus sur l'interaction forte.

Avant d'aborder la description précise de la diffusion X ou neutronique, signalons que d'autres sondes peuvent être utilisées. Par exemple, l'exploration de la surface d'un matériau se fait couramment par diffusion d'électrons lents en incidence rasante. D'une façon générale, toute particule massive peut être utilisée en fonction des besoins, et en ajustant son énergie de sorte que sa longueur d'onde de de Broglie ait la valeur appropriée.

Section efficace d'un réseau Quelle que soit la sonde utilisée, l'effet spectaculaire de l'ordre à longue distance présent dans un cristal est de produire une *hypersélectivité* en direction de la diffusion : on va voir que le cristal étant orienté d'une certaine façon par rapport à la direction du faisceau incident, le signal diffusé est nul presque partout, sauf pour des directions remarquables, la largeur angulaire de celles-ci étant réellement un infiniment petit. Cette spécificité de la diffusion par une assemblée organisée d'un (très) grand nombre de cibles porte le nom de *diffraction de Bragg.*

Dans la suite, on s'appuie sur les résultats donnés dans le chapitre 26 et on se place dans le cas d'un cristal tridimensionnel. Afin de considérer une situation concrète, les calculs seront faits à propos des neutrons, ce qui permettra d'ailleurs d'en dire un peu plus sur l'interaction de ceux-ci avec la matière, l'occasion de le faire n'étant pas si courante. La transcription au cas de la diffusion des rayons X est un simple jeu d'écriture à effectuer dans le potentiel d'interaction projectiles - cibles.

On se souvient que la section efficace est très simplement reliée à l'amplitude de diffusion (voir l'égalité (26.16)), laquelle est donnée par (26.33) ; pour une seule cible localisée au point spécifié par le vecteur $\vec{R}$, le potentiel créé par cette cible est une certaine fonction $V(\vec{r})$ ne dépendant que de $\vec{r}-\vec{R}$, que l'on note $V_{\vec{R}}(\vec{r})$; avec ce glissement de notation (et en prenant la constante de couplage g égale à 1), l'égalité (26.33) pour l'amplitude diffusée par cette cible se transcrit comme suit :

$$f_{k_0}^{(\vec{R})}(\theta,\,\phi)=-\frac{1}{4\pi}\int_{\mathbb{R}^3}\mathrm{e}^{-\mathrm{i}k_0\vec{n}.\vec{r}\,'}\,U_{\vec{R}}(\vec{r}\,')\psi_{\vec{k}_0}(\vec{r}\,')\,\mathrm{d}^3r'\ . \tag{29.42}$$

Dans l'approximation de Born (très largement suffisante ici), notant $\vec{k}_\mathrm{i}$ et $\vec{k}_\mathrm{f}$ les vecteurs d'onde incident et diffusé, et revenant au potentiel $V_{\vec{R}}(\vec{r}) \stackrel{\text{déf}}{=} \frac{\hbar^2}{2M}U_{\vec{R}}(\vec{r})$ de la cible localisée en $\vec{R}$, on a (voir (26.38)), M étant la masse du neutron :

$$f_{k_\mathrm{i}}^{(\vec{R})}(\theta,\,\phi)=-\frac{M}{2\pi\hbar^2}\int_{\mathbb{R}^3}\mathrm{e}^{-\mathrm{i}\vec{k}_\mathrm{f}.\vec{r}\,'}\,V_{\vec{R}}(\vec{r}\,')\,\mathrm{e}^{\mathrm{i}\vec{k}_\mathrm{i}.\vec{r}\,'}\,\mathrm{d}^3r'\ ; \tag{29.43}$$

en prenant des notations plus concises, et plus naturelles dans le présent contexte, $|\vec{k}\,\rangle$ désignant simplement une onde plane, ceci se récrit :

$$\boxed{f^{(\vec{R})}(\vec{k}_\mathrm{i},\vec{k}_\mathrm{f})=-\frac{M}{2\pi\hbar^2}\int_{\mathbb{R}^3}\mathrm{e}^{-\mathrm{i}\vec{k}_\mathrm{f}.\vec{r}\,'}\,V_{\vec{R}}(\vec{r}\,')\,\mathrm{e}^{\mathrm{i}\vec{k}_\mathrm{i}.\vec{r}\,'}\,\mathrm{d}^3r'\equiv-\frac{M}{2\pi\hbar^2}\langle\vec{k}_\mathrm{f}|V_{\vec{R}}|\vec{k}_\mathrm{i}\rangle} \tag{29.44}$$

la section efficace différentielle de la diffusion $\vec{k}_\mathrm{i}\rightarrow\vec{k}_\mathrm{f}$ par la cible située en $\vec{R}$, $\sigma^{(\vec{R})}(\vec{k}_\mathrm{i},\vec{k}_\mathrm{f})$, étant toujours égale à $|f^{(\vec{R})}(\vec{k}_\mathrm{i},\vec{k}_\mathrm{f})|^2$.

Avant d'examiner ce que fait un ensemble de diffuseurs, il convient de préciser le potentiel d'une cible, $V_{\vec{R}}(\vec{r})$, qui est le même pour tous les diffuseurs du réseau de Bravais. Pour cette raison, on a :

$$V_{\vec{R}}(\vec{r}) = V(\vec{r} - \vec{R}) \ , \tag{29.45}$$

où la fonction $V(\vec{r})$ est *indépendante* de $\vec{R}$. Comme annoncé plus haut, le neutron interagit principalement par une interaction de contact en ce qui concerne l'interaction forte, et par des interactions dipolaires magnétiques.

En ce qui concerne le couplage de contact, sa forme usuelle est[36] :

$$V_{\mathrm{c}}(\vec{r}) \stackrel{\text{déf}}{=} \frac{2\pi\hbar^2}{M}\, b\, \delta(\vec{r}) \ , \tag{29.46}$$

et s'appelle *pseudo-potentiel de Fermi* ; b est par définition la *longueur de diffusion*[37], qui varie d'un noyau à l'autre (et donc en particulier d'un isotope à l'autre), et dépend également du spin nucléaire.

La forme (29.46) entraîne que, pour un diffuseur unique, l'amplitude de diffusion (29.44) est simplement $f^{(\vec{R})}(\vec{k}_{\mathrm{i}}, \vec{k}_{\mathrm{f}}) = -b$, de sorte que les sections efficaces différentielle et totale sont :

$$\boxed{\frac{\mathrm{d}\sigma^{(\vec{R})}}{\mathrm{d}\Omega} = b^2 \ , \qquad \sigma^{(\vec{R})} = 4\pi b^2} \tag{29.47}$$

b rentre dans ces expressions de la même façon que le rayon d'une sphère dure pour un problème classique de diffusion, à un facteur 4 près[38].

En ce qui concerne le couplage magnétique, notamment avec les spins des électrons, l'interaction est de la forme :

$$V_{\mathrm{magn}} = -g_{\mathrm{n}} \frac{|e|}{2M_{\mathrm{n}}}\, \vec{S}_{\mathrm{n}}.\vec{\mathcal{B}} \ , \tag{29.48}$$

où $g_{\mathrm{n}} \simeq -3,826$ est le facteur anormal du neutron (voir p. 853) et $\vec{\mathcal{B}}$ le champ magnétique produit par les électrons ; $\vec{S}_{\mathrm{n}}$ est le spin du neutron ($S = 1/2$).

Le champ $\vec{\mathcal{B}}$ a deux origines (voir [235] pour plus de détails) : d'une part le champ dipolaire créé par le moment magnétique (de spin) des électrons (masse m), $\vec{\mu}_i$, égal à $g_{\mathrm{e}}\frac{e}{2m}\vec{S}_i$ (voir p. 853), d'autre part le champ créé par les boucles de courant constituées par les électrons en orbite (vitesse $\vec{v}_i$). Au total :

$$\vec{\mathcal{B}}(\vec{r}) = \frac{\mu_0}{4\pi} \sum_i \left(\vec{\nabla} \times \frac{\vec{\mu}_i \times (\vec{r}_i - \vec{r})}{\|\vec{r}_i - \vec{r}\|^3} + \frac{e\vec{v}_i \times (\vec{r}_i - \vec{r})}{\|\vec{r}_i - \vec{r}\|^3} \right) \ . \tag{29.49}$$

---

[36] Avec (29.45), l'interaction de contact avec le noyau situé en $\vec{R}$ est donc $\frac{2\pi\hbar^2}{M}\, b\, \delta(\vec{r} - \vec{R})$.

[37] Dans $\mathbb{R}^3$, la fonction $\delta(\vec{r})$ est homogène à L^{-3} ; il en résulte que b a bien la dimension d'une longueur. La terminologie *longueur de diffusion* est devenue habituelle dans d'autres contextes quand un potentiel de contact intervient ; on l'emploie par exemple dans le cas des réseaux optiques utilisés pour piéger des atomes froids [237] (voir aussi problème 30.7.14).

[38] Cette différence a fait l'objet d'un commentaire p. 1209.

Le terme d'interaction est donc $V_{\text{magn}} = V_{\text{dd}} + V_{\text{so}}$, avec :

$$\boxed{V_{\text{dd}}(\{\vec{r}_i\}, \vec{r}) \stackrel{\text{déf}}{=} g_{\text{e}} g_{\text{n}} \frac{e'^2}{4mM_{\text{n}}c^2} \sum_i \vec{S}_{\text{n}}.\vec{\nabla} \times \frac{\vec{S}_i \times (\vec{r}_i - \vec{r})}{\|\vec{r}_i - \vec{r}\|^3}} \tag{29.50}$$

et :

$$\boxed{V_{\text{so}}(\{\vec{r}_i\}, \vec{r}) \stackrel{\text{déf}}{=} g_{\text{n}} \frac{e'^2}{4mM_{\text{n}}c^2} \sum_i \vec{S}_{\text{n}} \cdot \Big(\vec{p}_i \times \frac{\vec{r}_i - \vec{r}}{\|\vec{r}_i - \vec{r}\|^3} - \frac{\vec{r}_i - \vec{r}}{\|\vec{r}_i - \vec{r}\|^3} \times \vec{p}_i\Big)} \tag{29.51}$$

V_{dd} est l'interaction spin - spin (dipôle - dipôle) – elle coïncide avec l'expression (27.26) où $(g_{\text{p}}, M_{\text{p}})$ et $(g_{\text{n}}, M_{\text{n}})$ sont échangés ; V_{so} est un couplage du genre spin - orbite, convenablement symétrisé pour tenir compte de la non-commutation des variables électroniques $\vec{r}_i$ et $\vec{p}_i$. Les deux termes V_{dd} et V_{so} sont du même ordre de grandeur, à savoir $\frac{E_{\text{I}}^2}{M_{\text{n}}c^2}$, soit environ 2×10^{-7} eV ; ils sont donc, tout naturellement, comparables à un couplage hyperfin. Globalement, l'interaction entre un neutron et les électrons d'un atome est de la forme :

$$V_{\text{magn}}(\{\vec{r}_i\}, \vec{r}) = \vec{S}_{\text{n}} . \vec{\mathcal{M}}(\vec{r}\,; \{\vec{r}_i\}, \{\vec{p}_i\}, \{\vec{S}_i\}) \ . \tag{29.52}$$

Compte tenu de l'implication des variables électroniques dans V_{magn}, la diffusion correspondante des neutrons donne aussi des informations sur la structure électronique dans les atomes et les molécules. L'interaction effective $V_{\vec{R}}(\vec{r})$ entre un neutron de position $\vec{r}$ et un atome (ou une molécule), centré en $\vec{R}$ et dans l'état $|\Psi_{\text{at}}\rangle$, s'obtient en sommant sur les degrés de liberté électroniques ; formellement :

$$V_{\text{magn eff}}(\vec{r}) = \vec{S}_{\text{n}} . \langle \Psi_{\text{at}} | \vec{\mathcal{M}}(\vec{r}\,; \{\vec{r}_i\}, \{\vec{p}_i\}, \{\vec{S}_i\}) | \Psi_{\text{at}} \rangle \ ; \tag{29.53}$$

c'est cette interaction qu'il faut entrer dans l'expression (29.43) à la place de $V(\vec{r})$ pour avoir l'amplitude de diffusion dans la première approximation de Born. Pour un atome dans un état singulet, seule l'interaction de type V_{so} survit à la moyenne avec $|\Psi_{\text{at}}\rangle$, auquel cas :

$$V_{\text{magn eff}}(\vec{r}) = \frac{g_{\text{n}} e'^2}{4mM_{\text{n}}c^2} \vec{S}_{\text{n}} . \langle \Psi_{\text{at}} | \sum_i \Big(\vec{p}_i \times \frac{\vec{r}_i - \vec{r}}{\|\vec{r}_i - \vec{r}\|^3} - \frac{\vec{r}_i - \vec{r}}{\|\vec{r}_i - \vec{r}\|^3} \times \vec{p}_i\Big) | \Psi_{\text{at}} \rangle \ . \tag{29.54}$$

L'interaction entre un neutron et une cible microscopique étant maintenant précisée, on peut en revenir à l'expression (29.44) pour prendre en compte non plus un mais les N diffuseurs constituant un échantillon macroscopique de volume L^3, en considérant d'abord le cas où la maille élémentaire ne contient qu'un seul diffuseur. Il suffit en fait de sommer sur $\vec{R} \in \mathcal{B}$ l'amplitude $f^{(\vec{R})}(\vec{k}_{\text{i}}, \vec{k}_{\text{f}})$ donnée en (29.44) pour obtenir celle de l'onde diffusée par le cristal :

$$f(\vec{k}_{\text{i}}, \vec{k}_{\text{f}}) = -\frac{M}{2\pi\hbar^2} \sum_{\vec{R} \in \mathcal{B}} \int_{\mathbb{R}^3} V_{\vec{R}}(\vec{r}\,') \, \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{r}\,'} \, \mathrm{d}^3 r' \qquad (\vec{q} \stackrel{\text{déf}}{=} \vec{k}_{\text{f}} - \vec{k}_{\text{i}}) \ . \tag{29.55}$$

On a vu que $V_{\vec{R}}(\vec{r}) = V(\vec{r}-\vec{R})$; posant $\vec{r}\,'-\vec{R} = \vec{r}\,''$ pour la variable muette d'intégration, on obtient :

$$f(\vec{k}_\mathrm{i}, \vec{k}_\mathrm{f}) = -\frac{M}{2\pi\hbar^2} \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}} \int_{\mathbb{R}^3} V(\vec{r}\,'')\,\mathrm{e}^{-\mathrm{i}\vec{q}.\vec{r}\,''}\,\mathrm{d}^3r'' \ . \qquad (29.56)$$

L'intégrale est indépendante de $\vec{R}$, et représente finalement la contribution du diffuseur de la cible située à l'origine (lequel n'a rien de particulier, en dehors de cette position conventionnelle). Il en résulte que l'amplitude diffusée par le cristal est de la forme :

$$\boxed{f(\vec{k}_\mathrm{i}, \vec{k}_\mathrm{f}) = \tilde{F}(\vec{q})\,S(\vec{q})} \qquad (29.57)$$

avec :

$$\boxed{S(\vec{q}) \stackrel{\text{déf}}{=} \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}} \ , \qquad \tilde{F}(\vec{q}) \stackrel{\text{déf}}{=} -\frac{M}{2\pi\hbar^2} \int_{\mathbb{R}^3} V(\vec{r})\,\mathrm{e}^{-\mathrm{i}\vec{q}.\vec{r}}\,\mathrm{d}^3r} \qquad (29.58)$$

$\tilde{F}(\vec{q})$, tout comme l'amplitude[39] $f(\vec{k}_\mathrm{i}, \vec{k}_\mathrm{f})$, est homogène à une longueur, et incorpore – *via* sa transformée de Fourier à $\vec{q}$ – les détails microscopiques d'un diffuseur quelconque. Quant à la somme $S(\vec{q})$, elle est en *facteur* de l'amplitude totale, et prend en compte à elle seule l'existence du réseau : c'est elle qui apporte l'*hypersélectivité* en direction, caractéristique de la diffusion coopérative, dite *diffraction de Bragg*, comme on va le voir.

Conditions de von Laue et de Bragg Le lien entre hypersélectivité et effets coopératifs de diffuseurs ordonnés peut maintenant s'établir facilement, et se tient en effet dans la somme $S(\vec{q})$ définie ci-dessus, qui incorpore tous les déphasages associés aux diffuseurs, comme le confirme l'argument suivant, d'inspiration purement classique.

Soit deux diffuseurs quelconques situés en deux points A et B (voir figure 29.3). La différence de chemin pour les deux ondes diffusées est IA+AJ, soit $R\cos\theta + R\cos\theta'$, ce que l'on peut écrire $\vec{R}.(-\vec{n}_\mathrm{i}) + \vec{R}.\vec{n}_\mathrm{f} = \vec{R}.(\vec{n}_\mathrm{f} - \vec{n}_\mathrm{i})$, où $\vec{R} = \vec{AB})$; pour deux diffuseurs séparés par $\vec{R}$ le déphasage est donc :

$$\frac{2\pi}{\lambda}\,\vec{R}.(\vec{n}_\mathrm{f} - \vec{n}_\mathrm{i}) \equiv \vec{R}.(\vec{k}_\mathrm{f} - \vec{k}_\mathrm{i}) \equiv \vec{R}.\vec{q} \ , \qquad \lambda = \frac{2\pi}{\|\vec{k}_\mathrm{i}\|} = \frac{2\pi}{\|\vec{k}_\mathrm{f}\|} \ . \qquad (29.59)$$

Classiquement, l'amplitude de l'onde diffusée par un réseau de Bravais $\mathcal{B}$ est ainsi de la forme[40] :

$$\Re A_0 \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}(\omega t - \vec{q}.\vec{R})} = \Re A_0 \mathrm{e}^{\mathrm{i}\omega t} \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}} \ , \qquad (29.60)$$

où l'on retrouve la somme $S(\vec{q})$ apparue spontanément dans le traitement quantique.

[39] Voir la définition de l'amplitude, éq. (26.10).
[40] $\Re$ désigne la partie réelle.

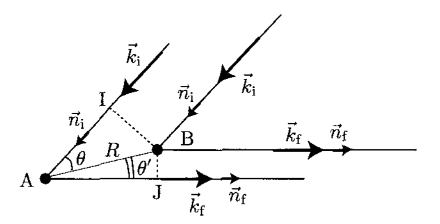


Figure 29.3: Déphasage pour deux diffuseurs A et B séparés du vecteur $\vec{R}$.

Calculons maintenant $S(\vec{q})$; la sommation sur $\vec{R}$ s'explicite comme suit :

$$S(\vec{q}) \stackrel{\text{déf}}{=} \sum_{\vec{R}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}} = \sum_{\{n_i\}\in\mathbb{Z}^3} \prod_{i=1}^{3} \mathrm{e}^{-\mathrm{i}n_i\vec{q}.\vec{a}_i} = \prod_{i=1}^{3} \sum_{n_i\in\mathbb{Z}} \mathrm{e}^{-\mathrm{i}n_i\vec{q}.\vec{a}_i} \ ; \tag{29.61}$$

en utilisant trois fois la définition du peigne de Dirac Ш(x) :

$$\boxed{\sum_{n=-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}nx} = 2\pi \sum_{p=-\infty}^{+\infty} \delta(x - p\,2\pi) \equiv \sum_{p=-\infty}^{+\infty} \delta(\frac{x}{2\pi} - p) \stackrel{\text{déf}}{=} Ш(x)} \tag{29.62}$$

on obtient :

$$S(\vec{q}) = \prod_{i=1}^{3} \Big[2\pi \sum_{p_i\in\mathbb{Z}} \delta(\vec{q}.\vec{a}_i - p_i\,2\pi)\Big] = \prod_{i=1}^{3} \Big[\sum_{p_i\in\mathbb{Z}} \delta(\frac{\vec{q}.\vec{a}_i}{2\pi} - p_i)\Big] \equiv \prod_{i=1}^{3} Ш(\frac{\vec{q}.\vec{a}_i}{2\pi}) \ . \tag{29.63}$$

L'apparition des peignes de Dirac révèle l'*hypersélectivité* en direction de la diffusion coopérative d'une infinité dénombrable d'objets *ordonnés* suivant un réseau de Bravais. Le transfert de moment $\vec{q}$ étant un vecteur de l'espace réciproque[41], il est naturel de le représenter (décomposer) sur des vecteurs primitifs $\vec{b}_j$ du réseau réciproque, en posant $\vec{q} = \sum_j \mu_j \vec{b}_j$; alors $\vec{q}.\vec{a}_i = 2\pi\,\mu_i$ en vertu de (29.38), et la somme $S(\vec{q})$ est :

$$S(\vec{q}) = \prod_{i=1}^{3} \Big[\sum_{p_i=-\infty}^{+\infty} \delta(\mu_i - p_i)\Big] \ ; \tag{29.64}$$

Ceci montre que la somme $S(\vec{q})$ est nulle si les μ_i ne sont pas *tous* des entiers. Inversement, si l'amplitude n'est pas nulle c'est parce que $\vec{q}$ est en fait un vecteur du réseau réciproque $\tilde{\mathcal{B}}$ – d'où la conclusion remarquable[42] :

$$\boxed{f(\vec{k}_\mathrm{i}, \vec{k}_\mathrm{f}) \neq 0 \implies \exists\ \vec{K} \in \tilde{\mathcal{B}}\ :\ \vec{k}_\mathrm{f} - \vec{k}_\mathrm{i} = \vec{K}} \tag{29.65}$$

[41] Il s'agit bien de l'*espace* réciproque, pas du réseau réciproque $\tilde{\mathcal{B}}$, où celui-ci est plongé. Les composantes de $\vec{q}$ sont pour l'instant des nombres quelconques, pas entiers.

[42] On remarquera que la double flèche n'est pas à double sens : une fois le réseau décoré avec son motif, il peut arriver que la condition de von Laue est satisfaite mais que la tache correspondante est en fait d'intensité *nulle* (voir plus loin le rôle du facteur de structure).

Cette condition, dite *condition de von Laue*[43], a une traduction géométrique simple ; compte tenu du fait que les deux vecteurs d'onde $\vec{k}_\mathrm{i}$ et $\vec{k}_\mathrm{f}$ ont même module, le vecteur formant la base d'un triangle isocèle construit sur $\vec{k}_\mathrm{i}$ et $\vec{k}_\mathrm{f}$ est un vecteur $\vec{K}$ du réseau réciproque. On parle de diffusion *cohérente* puisque l'intensité de diffraction est donnée par le module au carré de la *somme* $S(\vec{q})$, non par la somme des modules carrés de chacun de ses termes (voir Remarque 1 p. 1354).

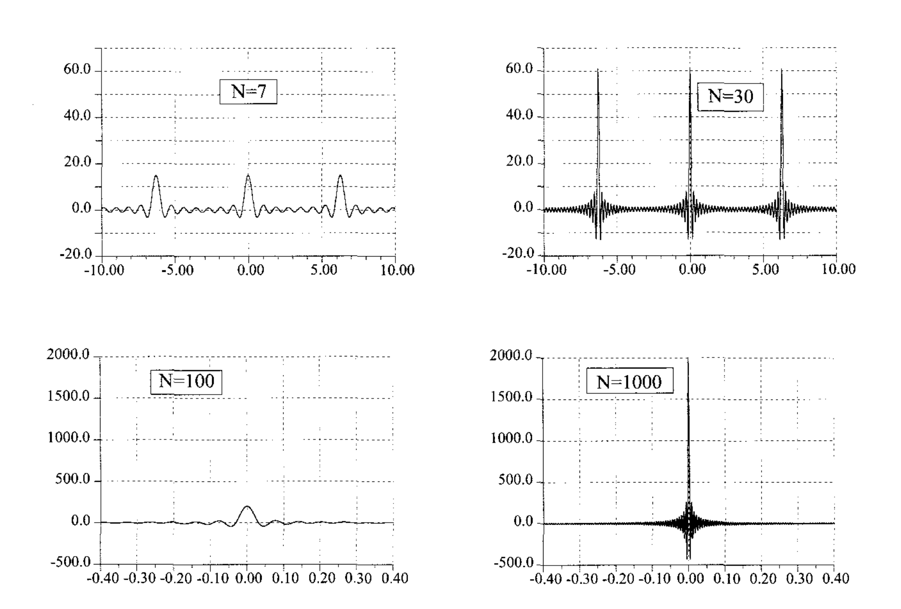


Figure 29.4: Variations de $S_i^{(N)}(\phi)$ en fonction de ϕ, (29.66), pour $N = 7$ et $N = 30$ d'une part, $N = 100$ et $N = 1000$ d'autre part (noter que les échelles sont les mêmes sur les deux couples de graphes). Imaginer $S_i^{(N)}$ quand $N \sim 10^{23}$...

♦ *Remarque*

Avant de continuer, justifions physiquement l'introduction des fonctions de Dirac, en raisonnant pour simplifier avec la seule direction d'espace définie par $\vec{a}_i$, et en supposant qu'il y a $2N+1$ diffuseurs le long de cette direction, placés symétriquement autour de l'origine. Dans ces conditions, la série $S_i(\vec{q})$ correspondante est en réalité une somme géométrique de $2N+1$ termes, $S_i^{(N)}$, qui se calcule facilement :

$$S_i^{(N)}(\phi) \stackrel{\text{déf}}{=} \sum_{n=-N}^{+N} \mathrm{e}^{in\phi} = \frac{\sin(2N+1)\frac{\phi}{2}}{\sin\frac{\phi}{2}} \quad , \quad \phi \stackrel{\text{déf}}{=} \vec{q}.\vec{a}_i \qquad (29.66)$$

Cette somme est 2π-périodique en ϕ ; le maximum est en $\phi = 0\ (2\pi)$, il vaut $2N+1$ et la largeur à mi-hauteur est de l'ordre de $\frac{2\pi}{2N+1}$; il est bien clair qu'avec

[43] On voit sans difficulté que cette condition est équivalente à la condition de Bragg $n\lambda = 2d\sin\theta$.

un nombre de diffuseurs de l'ordre du nombre d'Avogadro, remplacer la somme finie par un peigne de Dirac est pour le moins justifié (voir fig. 29.4). Noter que l'amplitude au maximum est proportionnelle au nombre de diffuseurs : l'intensité est donc proportionnelle au *carré* de ce nombre. ♦

29.3.3 Facteurs de structure (géométrique et atomique)

Afin de mettre en évidence l'hypersélectivité dans le contexte le plus simple, on a supposé que la maille primitive ne contenait qu'un seul diffuseur. Il convient de généraliser les résultats au cas d'un réseau décoré avec un motif quelconque, mais bien déterminé, ce qui correspond d'ailleurs à la situation la plus usuelle.

Le motif est complètement défini en stipulant la nature et la position $\vec{d_i}$ de ses n diffuseurs, relativement à une maille primitive (voir fig. 29.5). Le potentiel $V(\vec{r})$ défini au second membre de (29.45) est égal à la somme des potentiels créés par chaque diffuseur :

$$V(\vec{r}) = \sum_{i=1}^{n} V^{(i)}(\vec{r} - \vec{d_i}) \ , \tag{29.67}$$

de sorte que l'amplitude (29.56) a maintenant l'expression :

$$f(\vec{k}_{\mathrm{i}}, \vec{k}_{\mathrm{f}}) = -\frac{M}{2\pi\hbar^2} \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}} \int_{\mathbb{R}^3} \sum_{i=1}^{n} V^{(i)}(\vec{r} - \vec{d_i})\, \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{r}}\, \mathrm{d}^3 r \ . \tag{29.68}$$

Posant $\vec{r} - \vec{d_i} = \vec{r}\,'$, on obtient :

$$f(\vec{k}_{\mathrm{i}}, \vec{k}_{\mathrm{f}}) = \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}} \sum_{i=1}^{n} \tilde{F}_i(\vec{q})\mathrm{e}^{-\mathrm{i}\vec{q}.\vec{d_i}} \ . \tag{29.69}$$

où :

$$\tilde{F}_i(\vec{q}) \stackrel{\text{déf}}{=} -\frac{M}{2\pi\hbar^2} \int_{\mathbb{R}^3} V^{(i)}(\vec{r})\, \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{r}}\, \mathrm{d}^3 r \ , \tag{29.70}$$

contient la structure précise de l'élément i du motif. Par définition d'une maille élémentaire et du motif, la somme de droite dans (29.69) est indépendante de $\vec{R}$, d'où :

$$\boxed{f(\vec{k}_{\mathrm{i}}, \vec{k}_{\mathrm{f}}) = S(\vec{q}) \sum_{i=1}^{n} \tilde{F}_i(\vec{q})\mathrm{e}^{-\mathrm{i}\vec{q}.\vec{d_i}} \equiv S(\vec{q})\, F(\vec{q})} \tag{29.71}$$

On retrouve la factorisation de la somme $S(\vec{q})$ pour un réseau sans motif : l'hypersélectivité est encore la règle, toujours traduite par la condition de Bragg, mais l'intensité de chaque raie satisfaisant cette condition est maintenant en outre modulée par le second facteur, appelé *facteur de structure*. Il incorpore la position des éléments du motif par les vecteurs $\vec{d_i}$, et leur nature physique, le potentiel $V^{(i)}$ différant *a priori* d'un diffuseur à l'autre. Cette modulation de l'intensité des raies peut aller jusqu'à l'extinction pure

et simple de certaines d'entre elles. Les $\tilde{F}_i(\vec{q})$ sont appelés *facteurs de forme* (ou *facteur de structure atomique*) et représentent l'identité physique précise de chaque diffuseur. Avec ces définitions, tous les facteurs introduits sont, tout comme l'amplitude $f(\vec{k}_\mathrm{i}, \vec{k}_\mathrm{f})$, homogènes à une longueur, leur module au carré faisant penser à une section efficace.

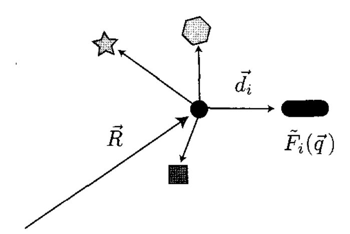


Figure 29.5: Notations utilisées dans (29.71).

On note que tous les facteurs de structure ou de forme sont toujours les transformés de Fourier d'une distribution de diffuseurs, discrète quand on considère des noyaux ou les positions des atomes, continue quand ce sont les électrons liés qui diffusent le rayonnement. Pour les rayons X, $\tilde{F}_i(\vec{q})$ est essentiellement la transformée de Fourier de la densité électronique.

♦ *Remarques*

1. *Diffusion cohérente* vs *diffusion incohérente*

   On a supposé tous les diffuseurs immobiles et organisés suivant un réseau ; l'hypersélectivité résulte immédiatement de ces deux hypothèses. Pour un fluide (gaz ou liquide), il n'existe pas d'ordre ou seulement un ordre à courte distance. L'amplitude de l'onde diffusée par l'échantillon est toujours proportionnelle à une somme d'amplitudes définissant le facteur $S(\vec{q})$ (29.58), mais la différence essentielle par rapport à un réseau (fixe) est que les positions des diffuseurs sont maintenant des grandeurs fluctuantes n'ayant que peu ou pas de corrélation les unes avec les autres. Pour un gaz, où le chaos moléculaire est la règle, tout comme pour un liquide où existe un ordre à *courte* portée, cet aléatoire fait passer de N^2 à N la dépendance de l'intensité par rapport au nombre de diffuseurs.

   Techniquement, le module au carré de $S(\vec{q})$ s'écrit :

$$|S(\vec{q})|^2 = \sum_{i,j=1}^{N} \mathrm{e}^{\mathrm{i}\vec{q}.(\vec{R}_i - \vec{R}_j)} = N\Big[1 + \sum_i \frac{1}{N}\sum_j (\mathrm{e}^{\mathrm{i}\vec{q}.(\vec{R}_i - \vec{R}_j)} - 1)\Big] \ . \qquad (29.72)$$

   Avec $N \gg 1$, la somme sur i porte sur la moyenne statistique $\langle \mathrm{e}^{\mathrm{i}\vec{q}.\vec{R}} - 1\rangle \stackrel{\text{déf}}{=} h(\vec{R})$, où $\vec{R}$ fixe les positions relatives de deux particules quelconques ; $h(\vec{R})$ est la fonction de corrélation (centrée) de paire. Pour un gaz, cette fonction est nulle partout, et le crochet vaut 1. Pour un liquide, $h(\vec{R})$ part de 0 en $\vec{R} = 0$, présente quelques

oscillations qui sont la signature de l'ordre local à courte portée, puis tend vers zéro aux grands $\|\vec{R}\|$; la somme sur i reconstitue alors $\rho \int_{\mathbb{R}^3} h(\vec{R})\mathrm{d}^3R$ où ρ est la densité du liquide[44].

Dans les deux cas, le crochet dans (29.72) est *intensif*, donnant au total une intensité proportionnelle à N, typique de la diffusion *incohérente*.

2. *Influence des vibrations de réseau*

   Dans le même ordre d'idées, mais revenant à un cristal, il est certain que si les atomes vibrent (un peu) autour d'une certaine position d'équilibre, la belle hypersélectivité risque d'en être affectée même si, au contraire d'un fluide, les positions *moyennes* continuent à suivre l'ordre à portée infinie caractéristique d'un cristal (parfait).

   Cette question se révèle en fait fort subtile : on peut montrer que, pour un cristal *tridimensionnel*, l'effet des vibrations de réseau est juste d'atténuer l'intensité de chaque raie de Bragg par un facteur appelé *facteur de Debye - Waller* (voir chapitre 31, section 31.7), mais n'anéantit pas la diffraction elle-même. En revanche, pour un cristal de dimensionnalité inférieure, les effets peuvent être beaucoup plus spectaculaires et conduire, notamment, à des largeurs (angulaires) traduisant les effets de taille finie de tout échantillon réel. Ce point sera rediscuté dans le chapitre 31, la quantification des vibrations d'un cristal faisant émerger tout naturellement des excitations élémentaires appelées *phonons*.

   Au passage, on retiendra l'idée importante suivante : des *défauts* ont souvent une importance mineure à haute dimension, mais se révèlent en général critiques pour les systèmes de basse dimensionnalité[45].

3. *Cas d'un cristal de dimension quelconque*

   La condition de von Laue $\vec{k}_\mathrm{i} = \vec{k}_\mathrm{f} + \vec{K}$ ressemble à une loi de conservation de l'impulsion (au facteur global $\hbar$ près). En raison de la symétrie de translation *discrète* du cristal, on peut dire que son impulsion est en fait définie seulement à $\hbar\vec{K}$ près, avec $\vec{K} \in \tilde{\mathcal{B}}$. C'est aussi pour cette raison que, pour un cristal *bidimensionnel*, la vraie condition de von Laue est $\vec{K} = \vec{K}_\parallel + \vec{K}_\perp$ (voir Remarque p. 1344). ♦

[44] C'est justement parce que l'ordre est à courte portée que cette intégrale est *finie*. Pour un cristal, elle est évidemment divergente, ce qui permet de comprendre intuitivement le changement de puissance de N à N^2 – le plus souvent, une divergence est le signal d'une modification de comportement.

[45] Cette idée est omniprésente en théorie des transitions de phase, et à propos des phénomènes critiques. Faire un trou dans un drap n'est pas si grave, couper un fil peut être catastrophique.

29.4 Exercices et problèmes

29.4.1 Molécule de van der Waals

L'interaction électrostatique dipôle-dipôle entre deux atomes ^{4}He a et b est décrite par le Hamiltonien :

$$H_{\rm dd} = \frac{1}{4\pi\varepsilon_0}\frac{1}{R^3}[\vec{d}_{\rm a}\cdot\vec{d}_{\rm b} - 3(\vec{d}_{\rm a}\cdot\vec{n})(\vec{d}_{\rm b}\cdot\vec{n})]$$

où $\vec{d}_{\rm a}$ et $\vec{d}_{\rm b}$ sont les opérateurs moment dipolaire électrique ; $\vec{R}$ est le vecteur reliant les deux noyaux, $\vec{n}$ le vecteur unitaire tel que $\vec{R} = R\,\vec{n}$.

Outre l'interaction dipôle-dipôle, il existe une forte répulsion à courte distance ($R \lesssim a_0$), qu'il n'est pas nécessaire de préciser ici. Dans la suite, on ne tient pas compte du spin.

Chaque atome est considéré comme un atome à deux niveaux[46], $|{\rm g}\rangle$ (fondamental), d'énergie $E_{\rm g}$, et $|{\rm e}\rangle$ d'énergie $E_{\rm e}$. $|{\rm gg}\rangle$ désigne l'état où les atomes sont chacun dans son état fondamental ; $|{\rm eg}\rangle$ est l'état où l'atome a est dans l'état excité $|{\rm e}\rangle$ et l'atome b dans l'état fondamental $|{\rm g}\rangle$. Il s'agit notamment de trouver les corrections d'énergie dues à $H_{\rm dd}$, puis de montrer l'existence de molécules géantes où seule la partie longue distance de l'interaction est pertinente.

État fondamental

1. Pour l'état fondamental, quelle est la correction d'énergie au premier ordre de la théorie des perturbations, $\Delta E^{(1)}$?

2. Écrire l'expression de la correction au second ordre, $\Delta E^{(2)}$ (avec $u = x, y, z$, on notera $d_{u\,{\rm ge}} \stackrel{\rm déf}{=} \langle {\rm g}|d_u|{\rm e}\rangle \in \mathbb{R}$). Représenter ce terme par un petit dessin *à la* Feynman.

État excités

Soit maintenant le cas où un atome se trouve dans l'état excité et l'autre dans l'état fondamental.

1. Trouver les énergies et les états propres dans la base $\{|{\rm eg}\rangle, |{\rm ge}\rangle\}$. Quel est le décalage ΔE^* par rapport à l'énergie $E_{\rm g} + E_{\rm e}$ en l'absence de $H_{\rm dd}$?

2. Comment la correction ΔE^* se différencie-t-elle de $\Delta E^{(2)}$? À quelle condition ce calcul est-il valide ?

[46] En pratique, l'état "fondamental" est en réalité un état excité métastable, par exemple le premier état triplet au-dessus du vrai fondamental.

3. À l'aide d'un schéma aussi précis que possible, représenter en fonction de R les deux branches de valeurs propres que fournit le traitement effectué ici.

Molécule de van der Waals

On considère maintenant deux états excités e′ et e″ ($E_{\mathrm{e}'} < E_{\mathrm{e}''}$) et les deux situations où un atome est dans l'état fondamental, et l'autre dans l'état e′ ou dans l'état e″ ; à l'ordre zéro, les états du couple d'atomes sont donc $|\mathrm{e}'\mathrm{g}\rangle$ (ou $|\mathrm{ge}'\rangle$), et $|\mathrm{e}''\mathrm{g}\rangle$ (ou $|\mathrm{ge}''\rangle$).

1. Sans faire aucun calcul et pour $R \gg a_0$, tracer sur une même figure l'énergie des deux branches associées aux états $(|\mathrm{e}'\mathrm{g}\rangle, |\mathrm{ge}'\rangle)$ et les deux branches associées aux états $(|\mathrm{e}''\mathrm{g}\rangle, |\mathrm{ge}''\rangle)$.
2. Expliquer comment, complétée par des couplages résiduels non explicités ici, cette approche permet de comprendre l'existence d'une "molécule de van der Waals", objet stable où seule la partie très longue distance de l'interaction entre atomes joue un rôle (molécule diatomique *géante*).
3. J. Léonard et al. [238] ont mis en évidence un couple (e, g) où la distance d'équilibre vaut à peu près $160\,a_0$, valeur pour laquelle la profondeur du puits est voisine de 2 GHz. Sachant que l'état vibrationnel fondamental est à environ 600 MHz au-dessus du minimum, estimer en eV/Å^2 la constante de force de la liaison de van der Waals entre les deux atomes d'hélium[47].

29.4.2 Gaz d'électrons dans le modèle du jellium

Le *jellium* est un modèle de métal simplifié dans lequel le réseau d'ions est remplacé par un fond continu homogène de charge positive. Pour le cas d'un alcalin, considéré dans toute la suite, le fond continu porte une charge totale $N|e|$, N étant le nombre d'électrons. Le système est supposé macroscopique ($N \ggg 1$) et on choisit des conditions aux limites cycliques (*à la* Born - von Kármán) définies relativement à une grande boîte cubique de côté L. Par ailleurs, tout effet de bord sera négligé. m et e désignent respectivement la masse et la charge d'un électron.

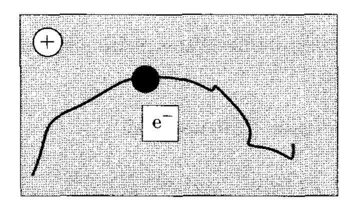


Figure 29.6: Représentation schématique du *jellium*.

[47]Pour en savoir plus, voir [239].

Il s'agit de calculer l'énergie de l'état fondamental du gaz d'électrons en traitant par perturbations au premier ordre la répulsion coulombienne entre électrons. Pour certains calculs, il est nécessaire techniquement de prendre une interaction écrantée, afin d'éviter les divergences dues à la portée infinie du potentiel coulombien ; ainsi, pour la répulsion entre deux électrons séparés par la distance r, on posera :

$$v(r) = \frac{e'^2}{r}\,\mathrm{e}^{-k_0 r} \quad . \tag{29.73}$$

où k_0^{-1} est la longueur d'écran. On prendra dès que possible la limite $k_0 \to 0$. Par ailleurs, on utilisera la substitution habituelle :

$$\sum_{\vec{k}} \quad \to \quad (\frac{L}{2\pi})^3 \int \mathrm{d}^3k \ . \tag{29.74}$$

Gaz parfait

Quand l'interaction de Coulomb est négligée, les N électrons forment un gaz parfait. Chaque électron est décrit par une spin-orbitale $\psi_{\vec{k}s}$ construite avec une onde plane :

$$\psi_{\vec{k}s}(\vec{r}) = L^{-3/2}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \otimes |s\rangle \qquad (s = \uparrow, \downarrow) \ . \tag{29.75}$$

1. Quelle est l'énergie $\varepsilon^{(0)}(\vec{k})$ d'un électron dans l'état (29.75) ?
2. Combien vaut le module k_F du vecteur de Fermi ? L'exprimer aussi en fonction de r_s défini comme :
$$\frac{4\pi}{3} r_s^3 = \frac{L^3}{N} \ . \tag{29.76}$$
3. Quelle est l'énergie cinétique du gaz d'électrons ?
4. La transformée de Fourier du potentiel coulombien écranté est :
$$\mathcal{V}(\vec{k}) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} \mathrm{d}^3r\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\,v(r) \ ; \tag{29.77}$$
calculer $\mathcal{V}(\vec{k})$.
5. La densité électronique varie-t-elle d'un point de l'espace à l'autre ? Quelle est la densité électronique totale ?
6. À l'aide de $\mathcal{V}(\vec{k} = 0)$, exprimer l'interaction :
   (a) entre les électrons et le fond continu positif,
   (b) du fond continu avec lui-même.
7. En déduire l'énergie totale du gaz parfait d'électrons immergé dans le fond continu positif.

Expression de la correction d'énergie au premier ordre

1. Soit $|\Psi_0\rangle$ l'état fondamental du gaz parfait d'électrons. Donner l'expression formelle de la correction d'énergie au premier ordre, ΔE, en fonction de la fonction v définie en (29.73) et de $|\Psi_0\rangle$.

2. L'état $|\Psi_0\rangle$ (normalisé) a pour expression :

$$|\Psi_0\rangle = \sqrt{N!}\,Y_{\rm A} \prod_{k\leq k_{\rm F},\, s} |\psi_{\vec{k}s}\rangle \equiv \sqrt{N!}\,Y_{\rm A}|\Phi_0\rangle \tag{29.78}$$

   où $Y_{\rm A}$ est le projecteur défini en (22.43), et $|\Phi_0\rangle$ un simple produit de spin-orbitales. Rappeler les principales propriétés de $Y_{\rm A}$.

3. Soit V l'opérateur associé à l'interaction coulombienne de tous les électrons. Montrer que :

$$\langle\Psi_0|V|\Psi_0\rangle = \langle\Phi_0|V\sum_{\lambda=1}^{N!}(-1)^{\sigma_\lambda}\,P_\lambda|\Phi_0\rangle \ . \tag{29.79}$$

4. En déduire que :

$$\langle\Psi_0|V|\Psi_0\rangle = \frac{1}{2}\sum_{\vec{k},\,k\leq k_{\rm F}}\ \sum_{\vec{k}',\,k'\leq k_{\rm F}}\ \sum_{s,\,s'}(\mathcal{J}_{\vec{k}\vec{k}'ss'} - \mathcal{K}_{\vec{k}\vec{k}'ss'}) \tag{29.80}$$

   où les termes directs et d'échange sont respectivement :

$$\mathcal{J}_{\vec{k}\vec{k}'ss'} = \langle\psi_{\vec{k}s}\psi_{\vec{k}'s'}|v|\psi_{\vec{k}s}\psi_{\vec{k}'s'}\rangle \ , \qquad \mathcal{K}_{\vec{k}\vec{k}'ss'} = \langle\psi_{\vec{k}s}\psi_{\vec{k}'s'}|v|\psi_{\vec{k}'s'}\psi_{\vec{k}s}\rangle \tag{29.81}$$

5. (a) Exprimer $\mathcal{J}_{\vec{k}\vec{k}'ss'}$ en fonction de L et $\mathcal{V}(0)$
   (b) En déduire que la contribution totale des termes directs compense très exactement les termes liés à l'existence du fond continu trouvés ci-dessus.

6. Exprimer $\mathcal{K}_{\vec{k}\vec{k}'ss'}$ en fonction de L, $\mathcal{V}(\vec{k}-\vec{k}')$ et un symbole de Kronecker.

7. En déduire que la correction au premier ordre de l'énergie du gaz d'électrons est donnée par :

$$\Delta E = -\frac{L^3}{(2\pi)^6}\int_{k\leq k_{\rm F}} {\rm d}^3k \int_{k\leq k_{\rm F}} {\rm d}^3k'\,\mathcal{V}(\vec{k}'-\vec{k}) \tag{29.82}$$

8. Montrer que tout se passe comme si un électron avait l'énergie :

$$\varepsilon(k) = \varepsilon^{(0)}(k) + \Sigma(k) \qquad (k\ \leq\ k_{\rm F}) \tag{29.83}$$

   Quelle est l'expression intégrale de la *self-energy* $\Sigma(k)$?

Interprétation des résultats

1. Montrer que, à la limite $k_0 = 0$, la *self-energy* $\Sigma(k)$ est donnée par :

$$\Sigma(k) = -\frac{1}{\pi}\, e'^2 k_\mathrm{F}\, S(\frac{k}{k_\mathrm{F}}) \tag{29.84}$$

   où la fonction $S(x)$ est :

$$S(x) \stackrel{\text{déf}}{=} \frac{1}{2} + \frac{1-x^2}{4x} \ln\left|\frac{x+1}{x-1}\right| \tag{29.85}$$

2. Tracer sommairement la fonction $S(x)$.
3. En définitive, l'énergie d'un électron peut se mettre sous la forme :

$$\varepsilon(k) = \varepsilon^{(0)}(k_\mathrm{F})\left[x^2 - C\frac{r_s}{a_0}S(x)\right] \qquad (x = k/k_\mathrm{F}) \tag{29.86}$$

   où $C \simeq 0,17$. Sur un même graphique, tracer $\varepsilon(k)/\varepsilon^{(0)}(k_\mathrm{F})$ et $\varepsilon^{(0)}(k)/\varepsilon^{(0)}(k_\mathrm{F})$ en fonction de x
4. Comparer la nouvelle largeur de bande à celle de la bande du gaz d'électrons sans interactions. De combien augmente-t-elle pour $r_s \sim 5\,a_0$?
5. On admet[48] que la vitesse moyenne d'un électron d'énergie $\varepsilon(k)$ définie en (29.83) est donnée par $\hbar^{-1}\frac{\mathrm{d}\varepsilon(k)}{\mathrm{d}k}$? Combien vaut-elle en $k = k_\mathrm{F}$? Qu'en pensez-vous ?
6. En revenant à l'expression intégrale de la *self-energy* obtenue plus haut, essayez d'identifier l'origine précise de cette anomalie

29.4.3 Le jellium : influence de la densité et de la portée des interactions sur les propriétés magnétiques de l'état fondamental

On reprend le modèle du *jellium* défini dans le problème 29.4.2, notant $|\vec{k}, s\rangle$ le *ket* associé à $\psi_{\vec{k}s}(\vec{r}) = L^{-3/2}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \otimes |s\rangle$ (voir (29.75)). $V(\vec{r})$ désigne l'interaction entre les électrons, traitée par perturbation au plus bas ordre.

Expression de l'énergie

L'état fondamental putatif est formé en mettant N_+ de spin $\uparrow$ et $N_- = N - N_+$ électrons de spin $\downarrow$ dans les états de basse énergie ; on introduit la polarisation $\zeta \stackrel{\text{déf}}{=} \frac{N_+ - N_-}{N}$.

[48] Ce résultat sera démontré dans le chapitre 30, sous-section 30.3.3.

1. Déterminer les rayons de Fermi $k_{F\pm}$ en fonction de r_s et de ζ.

2. Calculer la valeur moyenne de l'énergie cinétique par particule en fonction de ζ, r_s et $E_0 \stackrel{\text{déf}}{=} \frac{e'^2}{a_0}$.

3. Exprimer la correction d'énergie purement électronique à l'aide des termes directs et d'échange, puis en fonction de $\mathcal{V}(\vec{k}) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} V(\vec{r})\, e^{i\vec{k}.\vec{r}}$.

4. L'interaction fond-fond et électron-fond est supposée de même nature que l'interaction électron-électron. En déduire l'expression de la correction d'énergie totale au premier ordre :

$$\Delta E = -\frac{L^3}{2(2\pi)^6}\int d^3k_1 \int d^3k_2 \big[\theta(k_{F+}-k_1)\theta(k_{F+}-k_2)+\theta(k_{F-}-k_1)\theta(k_{F-}-k_2)\big]\,\mathcal{V}(\vec{k}_1-\vec{k}_2) \tag{29.87}$$

Choix du potentiel d'interaction

Interaction à longue portée Pour décrire cette situation, on choisit un potentiel de Yukawa $V(\vec{r}) \stackrel{\text{déf}}{=} \frac{q_1 q_2}{4\pi\varepsilon_0 r} e^{-\lambda r}$, dont on fera tendre la portée vers l'infini dès que possible.

1. Calculer explicitement ΔE en fonction de E_0, r_s et ζ.

2. En déduire l'énergie totale par particule, ε.

3. Considérant r_s comme un paramètre, étudier les variations de ε en fonction de ζ. Discuter la nature de l'état fondamental et montrer qu'il existe deux valeurs critiques r_{s1} et r_{s2} définies comme :

$$r_{s1} \stackrel{\text{déf}}{=} \frac{2}{5}(2^{1/3}+1)\Big(\frac{3\pi^2}{2}\Big)^{2/3} \simeq 5,45\ , \qquad r_{s2} \stackrel{\text{déf}}{=} \Big(\frac{3\pi^2}{2}\Big)^{2/3} \simeq 6,03\ . \tag{29.88}$$

   Qualifier les deux états caractérisés respectivement par $\zeta = 0$ et $\zeta = 1$.

4. Résumer ces résultats en traçant la variation du "paramètre d'ordre" ζ en fonction de r_s.

Interaction à courte portée Pour décrire la situation opposée, on choisit une interaction de Dirac, $V(\vec{r}) = g q_1 q_2 a_0^2\, \delta(\vec{r})$, où g est la constante de couplage sans dimension et $\delta(\vec{r}) = \delta(x)\delta(y)\delta(z)$.

Reprendre les questions 1 à 4 du paragraphe précédent ; montrer que les valeurs critiques sont ici :

$$r'_{s1} \stackrel{\text{déf}}{=} \frac{3g}{2\pi}\Big(\frac{2}{9\pi}\Big)^{2/3} \simeq 0,082\, g\ , \qquad r'_{s2} \stackrel{\text{déf}}{=} \frac{9g}{8\pi}\Big(\frac{4}{9\pi}\Big)^{2/3} \simeq 0,097\, g\ . \tag{29.89}$$

29.4.4 Diffusion de neutrons par un gaz diatomique

Un faisceau de neutrons de vecteur d'onde $\vec{k}_{\rm i}$ est diffusé élastiquement par un gaz diatomique enfermé dans un conteneur cubique de côté L. On désigne par $\vec{d}_i$ l'axe nucléaire de la molécule i $(1 \leq i \leq N)$, dont l'un des noyaux est repéré par le vecteur $\vec{R}_i$, par $\vec{q}$ le transfert de moment, et on se place dans l'approximation de Born.

1. Quelle est l'amplitude de diffusion pour la molécule i ?
2. En délaissant la diffusion vers l'avant, montrer que l'intensité de diffusion est proportionnelle à $N\langle|F(\vec{q})|^2\rangle$, où $F(\vec{q})$ est le facteur de structure d'une molécule, et où les $\langle ... \rangle$ désignent une moyenne sur les orientations de celle-ci.
3. En déduire la réponse du gaz en fonction de l'angle θ de diffusion.

physique

Chapitre 30

Électrons dans un cristal

La périodicité spatiale d'un cristal a des conséquences physiques d'une immense portée. Ce chapitre commence par la démonstration du théorème fondamental de Bloch, suivie des concepts qui s'en dégagent immédiatement (bande d'énergie, vitesse d'un électron, surface de Fermi). Sont ensuite exposés les deux grands schémas où la notion de gap *d'énergie émerge naturellement. L'explication de la distinction entre isolants et conducteurs est exposée, tout en attirant l'attention sur les limites de la théorie des bandes et des phénomènes importants dont elle ne peut rendre compte (isolant de Mott, instabilité de Peierls). Le chapitre se termine sur la discussion des* magnons*, superbe exemple de la richesse potentielle d'un modèle de première importance, la chaîne ferromagnétique de Heisenberg, belle occasion aussi de présenter l'*Ansatz *de Bethe dans un cadre assez simple.*

30.1 Préliminaires

Avant d'en venir à la description quantitative du mouvement d'une particule (un électron) dans un champ de force périodique dérivant d'un potentiel, il convient de donner quelques éléments justifiant la pertinence de cette problématique pour l'étude des solides cristallisés.

Un solide, ordonné ou non, est un système fort complexe, même si l'on s'en tient aux interactions élémentaires les plus simples, celle de Coulomb par exemple. En effet, un solide contient des noyaux et des électrons, une telle distinction n'étant pas d'ailleurs la plus naturelle : on a vu que, selon les cas et la nature physique du solide, il était plus

judicieux de le considérer comme un assemblage d'ions (solides ioniques ou métalliques) ou de molécules (solides moléculaires ou covalents).

Une fois effectuée la séparation de Born et Oppenheimer, la première étape[1] consiste à étudier le cristal sans vibrations, c'est-à-dire avec des *cœurs* fixes, et à ne considérer que les électrons de la couche externe qui sont aux premières loges pour encaisser les altérations survenant lors de la formation du solide (électrons de valence). Rappelons que cette distinction repose essentiellement sur un simple critère énergétique : les électrons de valence ont une énergie de liaison comparable, par nature, à celle d'une liaison chimique (quelques eV au plus), alors que les électrons internes sont très fortement liés (pour un élément de masse atomique même modeste, l'énergie de la couche (1s) est très vite de l'ordre du keV puisqu'elle croît en gros comme Z^2, partant de $13,6$ eV pour $Z = 1$).

À l'issue de ce dégrossissage, et en se plaçant toujours dans la première étape de l'approximation de Born et Oppenheimer, on reste *a priori* avec des ions fixes et les électrons de valence, ces derniers étant soumis au potentiel plus ou moins fort des premiers. En plus du potentiel périodique du réseau, les électrons sont évidemment en interaction mutuelle, au moins par leur répulsion électrostatique.

Comme on le sait, un tel problème est redoutable et n'est soluble qu'approximativement, à l'exception de quelques modèles simples se bornant à singer aussi bien que possible la répulsion électronique[2]. Dans une première approche, on peut décider d'oublier purement et simplement la répulsion de Coulomb entre électrons, ce qui, se souvenant du cas de l'hélium, est *a priori* très grossier. Une amélioration évidente consiste à prendre en compte partiellement la corrélation entre électrons par des approximations de champ moyen, du type Hartree - Fock par exemple, ce qui fait néanmoins rester dans le cadre d'une théorie à *une* particule.

De façon assez surprenante, nombreux sont les solides où les théories à une particule permettent d'expliquer un très grand nombre de propriétés (pas toutes, évidemment) et, en particulier, d'énoncer les critères permettant de savoir pourquoi il existe des isolants et des conducteurs. La justification précise de ce qui ressemble à un petit miracle n'est pas aisée, et prend comme point de départ la théorie phénoménologique des liquides normaux de Fermi que l'on doit à Landau ; cette théorie introduit fondamentalement la notion de *quasi-particule*, qui représente une particule nue enveloppée d'un *nuage* d'autres particules[3]. Si l'on retrouve la même idée de *renormalisation* d'un objet physique gardant une part de son intégrité comme pour l'atome *habillé* de Cohen-Tannoudji, la différence essentielle tient ici au fait qu'il s'agit authentiquement d'un effet à N-corps : le *rhabillage* d'une particule la métamorphosant en *quasi-particule* provient fondamentalement des interactions entre particules (nues) par ailleurs identiques, et constitue un exemple à la

[1] La deuxième étape consistera à étudier le mouvement des cœurs, qui donne lieu aux vibrations du solide et, dans l'hypothèse des petites oscillations, conduit à des oscillateurs harmoniques dont la quantification introduit naturellement les phonons du réseau – voir chapitre 31.

[2] L'un des plus célèbres est le modèle de Hubbard [240], dont la version sur réseau unidimensionnel (rigide) a été résolue en 1968 par Lieb et Wu [241], [242]. Si l'énoncé du problème est simple, sa résolution ne l'est pas...

[3] Pour reprendre l'image employée par Mattuck [181] : penser au *cow-boy* au galop sur son cheval dans le désert.

portée quasi-universelle d'effet *collectif*, distinct toutefois des *modes collectifs* également présents dans la plupart des systèmes condensés. Dans le langage de la Théorie quantique des champs, les quasi-particules sont des excitations de basse énergie au-dessus de l'état fondamental, comme toujours et par définition, *vide* de toute excitation.

Nozières et Luttinger ont démontré théoriquement la justesse et la pertinence de la théorie de Landau, tout en en délimitant le champ d'application ([243] - [245]). La justification théorique est très ardue, mais elle peut être schématisée à l'extrême, et mieux comprise, en se référant notamment à la notion commune de constante diélectrique. En raison de la forte densité d'électrons, des phénomènes d'écrantage viennent réduire considérablement la portée et l'importance de la répulsion de Coulomb[4] ; ceci se traduit en particulier par l'apparition dans la fonction de réponse du gaz d'électrons, d'une quantité analogue à une constante diélectrique, dont on sait classiquement qu'elle émerge naturellement dans une description effective à gros grain des interactions de Coulomb nues[5] dans un milieu traité comme un *continuum* : dans le cas présent, ce sont les autres électrons qui constituent le bain où se meuvent deux d'entre eux conventionnellement désignés. On peut aussi compléter ces idées intuitives en notant que les électrons disposent de surcroît d'un vaste espace (le solide) et sont, du fait de cette délocalisation, relativement éloignés les uns des autres. Pour toutes ces raisons, l'interaction *effective* est loin de ressembler à la répulsion de Coulomb nue et autorise, dans un grand nombre de cas, à délimiter un cadre de travail raisonnable préservant l'aspect *une* particule et avec lui toutes les commodités techniques qui s'y rattachent spécifiquement.

Ce cadre étant admis dans la suite, on reste ainsi avec quelques (quasi-)électrons pseudo-indépendants par atome, avec pour conséquence primordiale que la notion d'état à une particule garde toute sa pertinence, permettant de construire les états du solide en empilant les électrons disponibles dans ces états supposés prédéterminés, la fonction d'onde étant fondamentalement un déterminant de Slater. À ce stade de l'analyse, on voit que la toute première étape est de trouver le mouvement d'*un* électron dans un cristal supposé parfait, c'est-à-dire produisant un potentiel statique *périodique*, celui des ions immobiles situés aux nœuds d'un certain réseau de Bravais $\mathcal{B}$. Une fois admis le bien-fondé d'une théorie à une particule, c'est de cette symétrie spatiale, et d'elle seule, que vont découler les résultats importants établis dans la suite.

On en est ainsi réduit à devoir résoudre l'équation aux valeurs propres associée au Hamiltonien d'un unique électron dans un potentiel périodique :

$$\boxed{H = \frac{\vec{p}^{\,2}}{2m} + V(\vec{r}) \ , \qquad V(\vec{r}+\vec{R}) = V(\vec{r}) \quad \forall \vec{R} \in \mathcal{B}} \tag{30.1}$$

[4]On a vu que la densité électronique est globalement caractérisée par la longueur r_s (voir chapitre 29, éq. (29.22)), définissant le rayon de la sphère occupée en moyenne par un électron. Cette longueur est de quelques unités atomiques pour un métal – au contraire, à très faible densité, les électrons sont quasiment immobiles, au sens où l'énergie cinétique est très petite devant l'énergie potentielle (solide de Wigner).

[5]La constante diélectrique de l'eau est voisine de 80 : c'est le facteur de réduction de la force de Coulomb dans l'eau, comparé au vide. C'est bien pour cela que le sel (NaCl, un cristal ionique !) fond dans l'eau.

On devine que ceci n'est que le point de départ d'approximations d'une autre nature, puisque la classe de potentiels $V(\vec{r})$ exactement solubles est fort réduite[6]. Avant d'en venir à quelques schémas approximatifs de portée universelle (voir plus loin), il est important d'énoncer des résultats exacts, ne reposant que sur l'existence de la symétrie de translation discrète du réseau, traduite par l'invariance de l'énergie potentielle V dans toute translation d'un vecteur $\vec{R} \in \mathcal{B}$.

30.2 Théorème de Bloch

La symétrie de translation caractéristique de V entraîne que le Hamiltonien (30.1) commute avec tous les opérateurs de translation discrète $\mathcal{T}(\vec{R})$ associés chacun à une translation géométrique $T(\vec{R})$. De ce seul fait résulte le théorème de Bloch, dont l'énoncé est le suivant[7] :

Pour un Hamiltonien périodique du type (30.1), toute fonction (état) propre $\psi(\vec{r})$ peut être mise sous la forme d'une onde plane modulée en amplitude par une fonction $u(\vec{r})$ ayant la symétrie de périodicité du réseau :

$$\forall\psi(\vec{r}),\ H\psi(\vec{r}) = \varepsilon\psi(\vec{r}),\ \exists\, u(\vec{r}),\ \psi(\vec{r}) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}u(\vec{r}),\ u(\vec{r}) = u(\vec{r}+\vec{R}) \quad \forall\vec{R} \in \mathcal{B}\ . \tag{30.2}$$

Le point important est clairement[8] l'affirmation de la *périodicité* de la fonction $u(\vec{r})$. Bien sûr, chaque fonction propre et chaque énergie propre doivent être indexées par un (des) nombre(s) quantique(s), collectivement notés n, associés aux observables qui, avec H, forment un ECOC. En outre, l'apparition du vecteur $\vec{k}$ – paramètre homogène à un vecteur d'onde dont la signification sera donnée plus loin –, conduit, quand il le faut, à écrire (30.2) plus précisément (notations traditionnelles) $H\psi_{n\vec{k}}(\vec{r}) = \varepsilon_n(\vec{k})\,\psi_{n\vec{k}}(\vec{r})$ et :

$$\boxed{\psi_{n\vec{k}}(\vec{r}) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\,u_{n\vec{k}}(\vec{r}) \ , \quad u_{n\vec{k}}(\vec{r}+\vec{R}) = u_{n\vec{k}}(\vec{r}) \quad \forall\vec{R} \in \mathcal{B}} \tag{30.3}$$

L'affirmation centrale contenue dans le théorème de Bloch peut d'ailleurs être énoncée autrement. En effet, on a la suite d'égalités :

$$\psi_{n\vec{k}}(\vec{r}+\vec{R}) = \mathrm{e}^{\mathrm{i}\vec{k}.(\vec{r}+\vec{R})}\,u_{n\vec{k}}(\vec{r}+\vec{R}) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\,u_{n\vec{k}}(\vec{r}) \ , \tag{30.4}$$

[6] S'agissant d'exemple de potentiel périodique, celui de Kronig - Penney (un peigne de Dirac à une dimension) est un grand classique.

[7] Des arguments similaires peuvent se transposer pour des Hamiltoniens périodiques en temps ; le même théorème est alors connu en Mathématiques sous le nom de *Théorème de Floquet.*

Par ailleurs, si les paramètres de maille du réseau tendent vers zéro (potentiel constant !), (30.2) dit que les fonctions propres d'une particule libre sont des ondes planes $\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$: c'est l'invariance galiléenne.

[8] S'agissant de trouver n'importe quelle fonction propre de tout Hamiltonien, on peut toujours la *poser* sous la forme $\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\phi(\vec{r})$!

soit finalement :

$$\boxed{\psi_{n\vec{k}}(\vec{r}+\vec{R}\,) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\,\psi_{n\vec{k}}(\vec{r}\,) \quad \forall \vec{R} \in \mathcal{B}} \tag{30.5}$$

On utilisera selon les cas la forme (30.3) ou la forme (30.5).

Ce théorème peut être établi de plusieurs façons ; deux démonstrations sont proposées dans la suite, chacune d'entre elles ayant ses vertus et son utilité pour la référence ultérieure. Auparavant, il convient de souligner que les deux relations (30.3) et (30.5) reflètent visiblement une propriété de symétrie, ici de translation : le réseau se superpose à lui-même dans n'importe quelle translation de vecteur $\vec{R}$, de sorte que le module carré de la fonction d'onde doit être invariant :

$$|\psi_{n\vec{k}}(\vec{r}+\vec{R}\,)|^2 = |\psi_{n\vec{k}}(\vec{r}\,)|^2 \ . \tag{30.6}$$

Bien sûr – tout comme dans le cas de l'invariance de jauge (voir chapitre 17, section 17.4 et notamment éq. (17.149)) –, la fonction elle-même n'est pas invariante et acquiert au plus un facteur de phase sous l'opération de symétrie[9]. C'est ce que dit l'équation (30.5), tout en précisant quel est ici ce facteur de phase.

Première démonstration

Elle introduit explicitement les opérateurs de translation du réseau, $\mathcal{T}(\vec{R}\,)$, définis précisément comme suit[10] :

$$\boxed{\forall\, \phi(\vec{r}\,),\ \mathcal{T}(\vec{R}\,)\,\phi(\vec{r}\,) \overset{\text{déf}}{=} \phi(\vec{r}-\vec{R}\,)} \tag{30.8}$$

où l'on reconnaît la définition usuelle de l'opérateur agissant dans l'espace des états, associé à une transformation géométrique effectuée dans l'espace réel $\mathbb{R}^3$, ici la translation $T(\vec{R})$. Comme il s'agit d'opérations de symétrie, tous les $\mathcal{T}(\vec{R}\,)$ commutent avec H donné par (30.1) :

$$\forall \vec{R} \in \mathcal{B} \ : \quad [H, \mathcal{T}(\vec{R}\,)] = 0 \ ; \tag{30.9}$$

le sens physique de cette relation est clair : elle signifie qu'un état propre $\psi(\vec{r}\,)$ de H étant donné, son transformé $\mathcal{T}(\vec{R})\psi(\vec{r}\,)$ est aussi état propre avec la *même* énergie[11] (la moindre des choses). En effet, (30.9) dit que, pour toute fonction quelconque ϕ, on a

[9] Ici la phase est constante (elle est indépendante de la position $\vec{r}\,$) – dans le cas d'une transformation de jauge, au contraire, la phase est variable dans l'espace (voir éq. (17.149)). Elle garde toutefois toute sa pertinence : $\psi_{\vec{k}}$ et $\psi_{\vec{k}'}$ représentent deux états physiques distincts, tant que la différence $\vec{k}-\vec{k}'$ n'est pas un vecteur du réseau réciproque (voir plus loin).

[10] La définition (30.8) de la fonction transformée par l'opérateur de translation est la plus naturelle, et c'est elle qui est retenue pour la suite, en conformité avec le point de vue développé dans le chapitre 17 (voir (17.58)). On trouve parfois dans la littérature la définition "duale" alternative :

$$\mathcal{T}(\vec{R}\,)\,\phi(\vec{r}\,) = \phi(\vec{r}+\vec{R}\,) \tag{30.7}$$

[11] On retrouve ici la propriété fondamentale d'après laquelle un sous-espace dégénéré est *globalement* invariant dans les opérations de symétrie.

Pour un groupe de symétrie non abélien, les états propres de H ne sont pas forcément propres de tous les opérateurs de symétrie : penser au groupe des rotations dans $\mathbb{R}^3$ dans le cas d'un champ à

$H\mathcal{T}(\vec{R})\phi = \mathcal{T}(\vec{R})H\phi$; si de plus ϕ est un état propre de H, ψ, satisfaisant $H\psi = \varepsilon\psi$, alors[12] :

$$H(\mathcal{T}\psi) = \mathcal{T}H\phi = \mathcal{T}\varepsilon\psi = \varepsilon(\mathcal{T}\psi) \ ; \tag{30.10}$$

ceci montre que la fonction $\mathcal{T}\psi$ est propre avec la valeur propre ε, tout comme l'est ψ.

De toute évidence, et en s'appuyant sur la géométrie élémentaire, l'ensemble des $\mathcal{T}(\vec{R})$ peut être muni d'une structure de groupe, au moyen de la relation de définition qui s'impose :

$$\mathcal{T}(\vec{R})\,\mathcal{T}(\vec{R}\,') = \mathcal{T}(\vec{R}+\vec{R}\,') \ . \tag{30.11}$$

L'inverse est défini comme :

$$\mathcal{T}^{-1}(\vec{R}) = \mathcal{T}(-\vec{R}) \ . \tag{30.12}$$

Ce groupe a ici la puissance de $\mathbb{N}$ (translations *discrètes*) ; il est commutatif (abélien), puisque la somme $\vec{R}+\vec{R}\,'$ est commutative[13] :

$$\mathcal{T}(\vec{R})\mathcal{T}(\vec{R}\,') = \mathcal{T}(\vec{R}+\vec{R}\,') = \mathcal{T}(\vec{R}\,'+\vec{R}) = \mathcal{T}(\vec{R}\,')\mathcal{T}(\vec{R}) \Longleftrightarrow [\mathcal{T}(\vec{R}),\, \mathcal{T}(\vec{R}\,')] = 0 \tag{30.13}$$

Il en résulte que tous les $\mathcal{T}(\vec{R})$ ont des vecteurs propres communs. Par ailleurs, les $\mathcal{T}(\vec{R})$ sont des opérateurs unitaires, puisqu'ils forment un sous-ensemble discret des translations continues ; en conséquence, $\mathcal{T}^\dagger(\vec{R}) = \mathcal{T}^{-1}(\vec{R}) = \mathcal{T}(-\vec{R})$.

Comme H et $\mathcal{T}(\vec{R})$ commutent, ils ont des vecteurs propres en commun. Soit donc une fonction propre[14] $\psi(\vec{r})$ telle que les deux équations suivantes sont satisfaites :

$$H\,\psi(\vec{r}) = \varepsilon\,\psi(\vec{r}) \ , \tag{30.14}$$

$$\mathcal{T}(\vec{R})\,\psi(\vec{r}) = \tau(\vec{R})\,\psi(\vec{r}) \quad \forall \vec{R} \in \mathcal{B} \ , \tag{30.15}$$

où, tout naturellement, la valeur propre τ de $\mathcal{T}(\vec{R})$ dépend de $\vec{R}$. L'équation (30.14) n'est pas explicitement utile pour la suite : elle rappelle seulement qu'une fonction telle que ψ, commune à H et aux $\mathcal{T}(\vec{R})$, *existe.*

En conséquence de (30.15) et (30.11), on a :

$$\mathcal{T}(\vec{R})\,\mathcal{T}(\vec{R}\,')\,\psi(\vec{r}) = \tau(\vec{R})\,\tau(\vec{R}\,')\psi(\vec{r}) \tag{30.16}$$

et :

$$\mathcal{T}(\vec{R}+\vec{R}\,')\,\psi(\vec{r}) = \tau(\vec{R}+\vec{R}\,')\psi(\vec{r}) \ . \tag{30.17}$$

symétrie sphérique ; les harmoniques sphériques, propres de $\vec{L}^2$ et L_z, ne sont pas propres de tous les opérateurs de rotation ($L_\pm$ échange les Y_{lm} les unes dans les autres, et les L_u sont les générateurs des rotations). Toutefois, L_u ne modifie pas la valeur du nombre quantique l, puisque l'énergie ne change pas par rotation.

[12] les parenthèses sont redondantes algébriquement, mais sont utiles pour faire comprendre les choses.

[13] En outre, les générateurs des translations sont proportionnels à l'impulsion $\vec{p}$, qui est hermitique, et dont toutes les composantes commutent deux à deux. Pour une translation finie de vecteur $\vec{r}_0$, on a $\mathcal{T}(\vec{r}_0) = \mathrm{e}^{(1/\mathrm{i}\hbar)\vec{p}.\vec{r}_0}$. Le développement de l'exponentielle, avec $\vec{p} = -\mathrm{i}\hbar\vec{\nabla}$ montre que pour toute fonction entière $f(\vec{r})$ on a $\mathcal{T}(\vec{r}_0)f(\vec{r}) = f(\vec{r}-\vec{r}_0)$, en conformité avec (30.8).

[14] Pour la simplicité de l'écriture, on omet les indices $(n\,\vec{k})$ apparaissant dans les énoncés (30.3) et (30.5) du théorème de Bloch.

Les équations (30.16) et (30.17) sont vraies $\forall\,\psi$ également propre de H, donc pour une base complète. Il en résulte :

$$\tau(\vec{R}+\vec{R}\,') = \tau(\vec{R}\,)\,\tau(\vec{R}\,') \ . \tag{30.18}$$

τ est donc une fonction exponentielle – d'ailleurs, en vertu de l'unitarité des $\mathcal{T}$ toutes leurs valeurs propres sont de la forme $\mathrm{e}^{\mathrm{i}\times\text{phase réelle}}$, ce que l'on choisit d'écrire plus précisément, pour la commodité :

$$\tau(\vec{R}\,) = \mathrm{e}^{-\mathrm{i}\Phi(\vec{R}\,)} \ , \tag{30.19}$$

où Φ est une fonction à valeurs réelles ; (30.18) s'écrit alors :

$$\Phi(\vec{R}+\vec{R}\,') = \Phi(\vec{R}\,) + \Phi(\vec{R}\,') \ ; \tag{30.20}$$

cette relation dit que Φ est une forme linéaire de $\vec{R}$. Quel que soit le vecteur $\vec{R}$, une telle forme linéaire peut toujours s'écrire en produit scalaire[15] ; $\Phi(\vec{R}\,)$ est donc de la forme $\vec{k}.\vec{R}$, où $\vec{k}$ est quelconque et, en définitive :

$$\boxed{\tau(\vec{R}\,) = \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}} \qquad \vec{k} \in \mathbb{R}^D} \tag{30.21}$$

Compte tenu de ce résultat, la relation (30.15) s'écrit maintenant, $\forall\,\psi$ propre de H et des $\mathcal{T}(\vec{R})$:

$$\mathcal{T}(\vec{R}\,)\,\psi(\vec{r}\,) = \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}}\,\psi(\vec{r}\,) \ ; \tag{30.22}$$

d'après la définition (30.8), le premier membre est $\psi(\vec{r}-\vec{R}\,)$, d'où :

$$\psi(\vec{r}-\vec{R}\,) = \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}}\,\psi(\vec{r}\,) \ . \tag{30.23}$$

Ceci n'est autre que la formulation (30.5) du théorème de Bloch, avec le vecteur $-\vec{R}$.

Notons qu'il n'y a pas *équivalence* entre les deux équations (30.14) et (30.15) : la première entraîne la seconde, mais la réciproque est fausse :

$$H\,\psi(\vec{r}\,) = \varepsilon\,\psi(\vec{r}\,) \implies \mathcal{T}(\vec{R}\,)\,\psi(\vec{r}\,) = \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}}\,\psi(\vec{r}\,) \tag{30.24}$$

Par exemple, soit deux Hamiltoniens H et H', tous deux invariants par le *même* groupe de translations, mais ne commutant pas entre eux. Si $|\psi_{n\vec{k}}\rangle$ est propre de H et des $\mathcal{T}(\vec{R})$, il ne peut être aussi vecteur propre de H' ; on peut seulement affirmer que l'élément de matrice $\langle\psi_{n\vec{k}_1}|H'|\psi_{m\vec{k}_2}\rangle$ est nul si $\vec{k}_1 \neq \vec{k}_2$, autrement dit que H' couple exclusivement entre eux les états propres de H ayant tous le même vecteur d'onde $\vec{k}$.

[15] C'est l'une des définitions possibles de l'espace dual d'un espace vectoriel.

Deuxième démonstration

La démonstration précédente est en un sens cantonnée à l'espace réel. Au contraire, celle qui suit se place d'emblée dans l'espace réciproque, exploite les propriétés de la transformée de Fourier du potentiel périodique $V(\vec{r})$ et détermine en fait la fonction propre $\psi(\vec{r})$ par sa transformée de Fourier $\phi(\vec{q})$.

Ici arrive une petite difficulté technique : la fonction ψ est étendue sur tout le cristal (en principe infini) et donc n'a pas de transformée de Fourier au sens usuel. Pour contourner cette difficulté, on introduit provisoirement un cristal de taille finie, construit pour la commodité sur un (très grand) nombre de mailles primitives. L'espace direct étant dès lors *fini*, la transformée de Fourier de ψ est une intégrale sur un domaine fini (mais aussi grand que l'on veut), de sorte que la question de son existence ne se pose pas. Tout naturellement, on a envie de prendre des conditions aux limites franches, en disant que ψ s'annule strictement en-dehors de la grande boîte (l'électron ne peut sortir de la boîte) ; en fait, ces conditions se révèlent peu commodes, techniquement parlant. Pour cette raison, on choisit à nouveau les conditions *cycliques* de Born - von Kármán (voir p. 1121), partant de l'idée que, le système physique étant par hypothèse (implicite) bien conditionné, le détail de ce qui se passe à la frontière de la grande boîte est sans importance physique (les éventuelles différences entre un jeu de conditions aux limites et un autre se traduisent par des termes s'annulant dans la limite d'un système infiniment étendu – l'analogue ici de la *limite thermodynamique*).

Un jeu de vecteurs primitifs[16] de $\mathcal{B}$, $\vec{a}_i$ $(i = 1, 2, 3)$, étant effectivement choisi, ces conditions identifient le point[17] $\vec{r}$ et le point $\vec{r} + N_i\vec{a}_i$ pour chaque direction de $\mathcal{B}$ $(i = 1, 2, 3,\ N_i \gg 1)$. Ces deux points étant confondus (ce sont donc finalement les mêmes, l'espace étant replié sur lui-même), il en résulte :

$$f(\vec{r} + N_i\vec{a}_i) = f(\vec{r}) \quad \forall \text{ fonction } f \qquad (i = 1, 2, 3\,;\, N_i \gg 1)\ . \tag{30.26}$$

Ceci étant admis, le vecteur $\vec{k}$ apparaissant dans (30.5) est maintenant quantifié à toute petite échelle ; en effet, en prenant pour f une onde plane, on a $\mathrm{e}^{\mathrm{i}\vec{k}.(\vec{r}+N_i\vec{a}_i)} = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}$, soit :

$$\boxed{\mathrm{e}^{\mathrm{i}N_i\vec{k}.\vec{a}_i} = 1} \tag{30.27}$$

Comme on l'a vu à plusieurs reprises, cette égalité contraint de fait les valeurs de $\vec{k}$; ici, plus précisément, développant $\vec{k}$ sur des vecteurs primitifs $\vec{b}_j$ du réseau réciproque, $\vec{k} = \sum_j \kappa_j\vec{b}_j$, et tenant compte de (29.38), on voit que les composantes κ_j satisfont $\mathrm{e}^{2\pi\mathrm{i}N_j\kappa_j} = 1$, ce qui impose $\kappa_j =$ entier relatif $\times\ \frac{1}{N_j}$; le vecteur $\vec{k}$ admet donc la

[16] Pour simplifier, on suppose désormais que le réseau est tridimensionnel.

[17] Noter que la cyclicité impose aussi :

$$\mathcal{T}(\vec{R} = N_i\vec{a}_i) = \mathbf{1} \quad \forall i\ . \tag{30.25}$$

décomposition :

$$\boxed{\vec{k} = \sum_{j=1}^{3} \frac{m_j}{N_j}\vec{b}_j \qquad (m_j \in \mathbb{Z})} \tag{30.28}$$

Comme toujours, l'application de conditions aux limites produit la quantification (les composantes de $\vec{k}$ sont des multiples entiers de quantités finies bien définies) ; le vecteur d'onde $\vec{k}$ est maintenant quantifié suivant une grille de dimensions linéaires $\frac{1}{N_j}$; chaque volume élémentaire est un gros point (*cellule, grain*), image en miniature d'une maille primitive du réseau réciproque, et a pour volume $\delta^3 k$:

$$\delta^3 k = \frac{1}{N_1 N_2 N_3}\,|\vec{b}_1.(\vec{b}_2\times\vec{b}_3)| \equiv \frac{1}{N}\,|(\vec{b}_1,\,\vec{b}_2,\,\vec{b}_3)| = \frac{1}{N}\,\frac{(2\pi)^3}{v}\ , \tag{30.29}$$

v désignant toujours le volume d'une maille primitive de $\mathcal{B}$ (on a utilisé (29.39)).

Maintenant, toute fonction $\psi(\vec{r})$ satisfaisant ces conditions cycliques peut être développée sur des ondes planes satisfaisant elles-mêmes ces conditions. On peut donc écrire *a priori* :

$$\psi(\vec{r}) = \sum_{\vec{q}} \phi(\vec{q})\,\mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}}\ , \tag{30.30}$$

où $\phi(\vec{q})$ est la composante de Fourier de $\psi(\vec{r})$ pour la valeur $\vec{q}$ et où la sommation (discrète) court sur tous les $\vec{q}$ de la forme :

$$\vec{q} = \sum_{j=1}^{3} \frac{m_j}{N_j}\vec{b}_j \qquad (m_j \in \mathbb{Z}) \quad \Longleftrightarrow \quad \mathrm{e}^{\mathrm{i}N_i\vec{q}.\vec{a}_i} = 1\ . \tag{30.31}$$

m_j étant dans tout $\mathbb{Z}$, $\vec{q}$ n'est pas borné et décrit "tout" $\mathbb{R}^3$ (et pas seulement une maille de $\tilde{\mathcal{B}}$), cet espace étant simplement *maillé* par les petits volumes $\delta^3 k$ introduits par les conditions cycliques ; en quelque sorte, la somme $\sum_{\vec{q}}$ dans (30.30) est un précurseur de l'intégrale[18]. Il n'y a pas de doute que l'expression (30.30) fournit une fonction satisfaisant les conditions cycliques puisque, selon (30.31) :

$$\psi(\vec{r}+N_i\vec{a}_i) = \sum_{\vec{q}} \phi(\vec{q})\,\mathrm{e}^{\mathrm{i}\vec{q}.(\vec{r}+N_i\vec{a}_i)} = \sum_{\vec{q}} \phi(\vec{q})\,\mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}} = \psi(\vec{r})\ . \tag{30.32}$$

L'objectif final étant de trouver $\psi(\vec{r})$, il est équivalent, selon (30.30), de trouver ses composantes de Fourier $\phi(\vec{q})$. Pour cela, exploitons maintenant la périodicité de $V(\vec{r})$, la fonction $V(\vec{r})$ étant suffisamment régulière pour admettre le développement :

$$V(\vec{r}) = \sum_{\vec{K}\in\tilde{\mathcal{B}}} \mathcal{V}(\vec{K})\,\mathrm{e}^{-\mathrm{i}\vec{K}.\vec{r}}\ , \tag{30.33}$$

[18]Il ne lui manque que le facteur $\delta^3 k$ pour être une somme de Darboux.
À ce stade, il n'y a pas d'autre restriction pour les k_j, donc pas de restriction sur les m_j relativement aux N_j. On verra par la suite que deux états caractérisés par $\vec{k}$ et $\vec{k}+\vec{K}$ ($\vec{K}\in\tilde{\mathcal{B}}$) représentent en fait le même état physique, ce qui permettra d'épuiser tous les états physiques distincts en cantonnant $\vec{k}$ à une maille primitive de $\tilde{\mathcal{B}}$, soit en se bornant (par exemple) à $1 \leq m_j \leq N_j$.

où la somme court maintenant sur le réseau réciproque $\tilde{\mathcal{B}}$; comme $\mathrm{e}^{\mathrm{i}\vec{K}\cdot(\vec{r}+\vec{R})} = \mathrm{e}^{\mathrm{i}\vec{K}.\vec{r}}$, le second membre est bien l'expression d'une fonction périodique dans l'espace réel – ce n'est rien d'autre qu'une série de Fourier spatiale. La relation inverse permettant de trouver les composantes de Fourier $\mathcal{V}(\vec{K})$ connaissant $V(\vec{r})$ est :

$$\mathcal{V}(\vec{K}) = \frac{1}{v}\int_{\text{maille élémentaire}} V(\vec{r})\,\mathrm{e}^{\mathrm{i}\vec{K}.\vec{r}}\,\mathrm{d}^3r \ , \tag{30.34}$$

l'intégration étant restreinte à une maille élémentaire de $\mathcal{B}$. En effet, on a d'une part :

$$\int_{\text{maille élémentaire}} V(\vec{r})\,\mathrm{e}^{\mathrm{i}\vec{K}.\vec{r}}\,\mathrm{d}^3r = \sum_{\vec{K}'\in\tilde{\mathcal{B}}} \mathcal{V}(\vec{K})\int_{\text{maille élémentaire}} \mathrm{e}^{\mathrm{i}(\vec{K}-\vec{K}').\vec{r}}\,\mathrm{d}^3r \ . \tag{30.35}$$

D'autre part, $\vec{K}-\vec{K}'$ est de la forme $\sum_j m_j\vec{b}_j$ et $\vec{r} = \sum_i x_i\vec{a}_i$ avec $x_i \in [0,\,1]$ puisque $\vec{r}$ est restreint à une maille élémentaire ; l'intégrale à droite dans (30.35) est donc :

$$\int_{\text{maille élémentaire}} \mathrm{e}^{2\pi\mathrm{i}\sum_j m_j x_j}\,\mathrm{d}^3r = v\prod_{i=1}^{3}\int_0^1 \mathrm{e}^{2\pi\mathrm{i}m_j x_j}\,\mathrm{d}x_j \ . \tag{30.36}$$

Chaque facteur intégral vaut 0 si son m_j est non-nul ; au total, l'intégrale n'est différente de zéro que si tous les m_j sont nuls, et vaut alors v, d'où :

$$\boxed{\int_{\text{maille élémentaire}} \mathrm{e}^{\mathrm{i}(\vec{K}-\vec{K}').\vec{r}}\,\mathrm{d}^3r = v\,\delta_{\vec{K}\vec{K}'} \qquad (\vec{K},\,\vec{K}')\in\tilde{\mathcal{B}}^2} \tag{30.37}$$

où $\delta_{\vec{K}\vec{K}'}$ est le symbole de Kronecker[19] ; la relation inverse (30.34) découle alors immédiatement. Notons qu'avec la définition (30.34), $V(\vec{r})$ et $\mathcal{V}(\vec{K})$ ont la même dimension physique, celle d'une énergie.

Avant de continuer, quelques remarques donnant lieu à des simplifications dans la suite :

- Sauf cas exceptionnel[20], on peut toujours choisir

$$\boxed{\mathcal{V}(\vec{K}=0) = 0} \tag{30.39}$$

puisque $V(\vec{r})$ est défini à une constante additive près. Adopter cette relation, c'est fixer à zéro la moyenne spatiale sur une maille primitive du potentiel périodique (cette convention sera utile pour l'approximation des électrons presque libres ; au stade actuel, elle n'est pas nécessaire).

[19] Transposées aux variables espace - vecteur d'onde, les relations (30.33), (30.34) et (30.37) ne sont rien d'autre que $(n,\,n'\in\mathbb{Z})$:

$$f(t) = \sum_{n\in\mathbb{Z}} f_n\,\mathrm{e}^{-\mathrm{i}n\omega t} \ , \quad f_n = \frac{\omega}{2\pi}\int_{\text{période}} f_n\,\mathrm{e}^{\mathrm{i}n\omega t}\,\mathrm{d}t \ , \quad \int_{\text{période}} \mathrm{e}^{\mathrm{i}(n-n')\omega t}\,\mathrm{d}t = 2\pi\,\delta_{nn'} \ , \tag{30.38}$$

pour le couple énergie - pulsation dans le cas d'une fonction de période $T = \frac{2\pi}{\omega}$.

[20] Une exception notable est le potentiel de Kronig - Penney, voir problème 17.8.6.

- $V(\vec{r})$ étant une fonction à valeurs réelles, $\mathcal{V}(\vec{K})$ a la symétrie :

$$\mathcal{V}^*(\vec{K}) = \mathcal{V}(-\vec{K}) \ . \tag{30.40}$$

- Si le cristal a la symétrie d'inversion d'espace ($V(-\vec{r}) = V(\vec{r})$), alors $\mathcal{V}(\vec{K})$ est aussi réel puisque l'intégrale $\int_{\mathbb{R}^3} V(\vec{r}) \sin(\vec{K}.\vec{r})\, \mathrm{d}^3 r$, possédant un intégrand impair, est nulle. En pareil cas :

$$\mathcal{V}(-\vec{K}) = \mathcal{V}(\vec{K}) = \mathcal{V}(\vec{K})^* \ . \tag{30.41}$$

Revenant à l'équation aux valeurs propres[21] pour $\psi(\vec{r})$ en représentation-q :

$$\Big[-\frac{\hbar^2}{2m}\vec{\nabla}^{\,2} + V(\vec{r})\Big]\, \psi(\vec{r}) = \varepsilon\, \psi(\vec{r}) \ , \tag{30.42}$$

et y injectant les développements (30.30) et (30.33), on obtient deux types de termes ; d'une part :

$$-\frac{\hbar^2}{2m}\vec{\nabla}^{\,2} \sum_{\vec{q}} \phi(\vec{q})\, \mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}} = -\frac{\hbar^2}{2m} \sum_{\vec{q}} (\mathrm{i}\vec{q})^2 \phi(\vec{q}) \mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}} \ ; \tag{30.43}$$

d'autre part :

$$V(\vec{r}) \sum_{\vec{q}} \phi(\vec{q}) \mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}} = \sum_{\vec{q}} \sum_{\vec{K}} \mathcal{V}(\vec{K}) \phi(\vec{q}) \mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}} \mathrm{e}^{-\mathrm{i}\vec{K}.\vec{r}} = \sum_{\vec{q}\,'} \sum_{\vec{K}\in\tilde{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q}\,' + \vec{K}) \mathrm{e}^{\mathrm{i}\vec{q}\,'.\vec{r}} \ . \tag{30.44}$$

En identifiant maintenant dans les deux membres de l'équation aux valeurs propres (30.42) le coefficient de $\mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}}$, on obtient :

$$+\frac{\hbar^2 \vec{q}^{\,2}}{2m}\, \phi(\vec{q}) + \sum_{\vec{K}\in\tilde{\mathcal{B}}} \mathcal{V}(\vec{K})\, \phi(\vec{q} + \vec{K}) = \varepsilon \phi(\vec{q}) \ ; \tag{30.45}$$

posant :

$$\varepsilon^{(0)}(\vec{q}) \stackrel{\text{déf}}{=} \frac{\hbar^2 \vec{q}^{\,2}}{2m} \ , \tag{30.46}$$

il vient finalement :

$$\boxed{\forall \vec{q} \ : \ [\varepsilon - \varepsilon^{(0)}(\vec{q})]\, \phi(\vec{q}) = \sum_{\vec{K}\in\tilde{\mathcal{B}}} \mathcal{V}(\vec{K}) \phi(\vec{q} + \vec{K})} \tag{30.47}$$

Par inspection, on voit que le système linéaire[22] (30.47) couple exclusivement entre elles les composantes $\phi(\vec{q})$ ne différant que d'un vecteur $\vec{K}$ appartenant à $\tilde{\mathcal{B}}$. Autrement dit, $\vec{q}$ étant confiné dans une maille primitive de $\tilde{\mathcal{B}}$, seuls les vecteurs de la forme $\vec{q} + \vec{K}$,

[21] Noter que la forme de l'équation (30.42) exclut la présence d'un champ magnétique.

[22] Il s'agit bien d'un *système* : il existe une équation pour chaque valeur de $\vec{q}$ choisie sur la grille plongée dans $\mathbb{R}^3$ définie plus haut.

$\vec{K}$ quelconque[23] dans $\tilde{\mathcal{B}}$, apparaissent dans la sommation de l'expression (30.30). En désignant par $\vec{k}$ une valeur particulière pour $\vec{q}$, on obtient une fonction ψ paramétrée par $\vec{k}$, que l'on note $\psi_{\vec{k}}(\vec{r})$ et dont le développement est de la forme :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{K}\in\tilde{\mathcal{B}}} \phi(\vec{k}+\vec{K})\,\mathrm{e}^{\mathrm{i}(\vec{k}+\vec{K}).\vec{r}} \equiv \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \sum_{\vec{K}\in\tilde{\mathcal{B}}} \phi(\vec{k}+\vec{K})\,\mathrm{e}^{\mathrm{i}\vec{K}.\vec{r}} \ . \tag{30.48}$$

Maintenant, la somme située le plus à droite dans (30.48) est visiblement une fonction de $\vec{r}$ ayant la symétrie de translation du réseau direct : si on remplace dans la sommation $\mathrm{e}^{\mathrm{i}\vec{K}.\vec{r}}$ par $\mathrm{e}^{\mathrm{i}\vec{K}.(\vec{R}+\vec{r})}$, rien ne change puisque $\mathrm{e}^{\mathrm{i}\vec{K}.\vec{R}}=1$: cette somme n'est rien d'autre que la fonction $u_{\vec{k}}$ introduite dans la formulation (30.3) du théorème de Bloch ; (30.48) peut ainsi se récrire, pour la commodité ultérieure :

$$\psi_{\vec{k}}(\vec{r}) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}u_{\vec{k}}(\vec{r}) \ , \quad u_{\vec{k}}(\vec{r}) \overset{\text{déf}}{=} \sum_{\vec{K}'\in\tilde{\mathcal{B}}} \phi(\vec{k}-\vec{K}')\,\mathrm{e}^{-\mathrm{i}\vec{K}'.\vec{r}} = u_{\vec{k}}(\vec{r}+\vec{R}) \quad \forall\,\vec{R}\in\mathcal{B} \ . \tag{30.49}$$

Ainsi s'achève la deuxième démonstration de ce théorème. Par ailleurs, l'égalité de droite montre bien que :

$$u_{\vec{k}+\vec{K}_0}(\vec{r}) = \sum_{\vec{K}'\in\tilde{\mathcal{B}}} \phi(\vec{k}+\vec{K}_0-\vec{K}')\,\mathrm{e}^{-\mathrm{i}\vec{K}'.\vec{r}} = \sum_{\vec{K}''\in\tilde{\mathcal{B}}} \phi(\vec{k}-\vec{K}'')\,\mathrm{e}^{-\mathrm{i}(\vec{K}''+\vec{K}_0).\vec{r}} = \mathrm{e}^{-\mathrm{i}\vec{K}_0.\vec{r}}u_{\vec{k}}(\vec{r}) \quad \forall\,\vec{K}_0\in\tilde{\mathcal{B}} \ , \tag{30.50}$$

d'où :

$$\psi_{\vec{k}+\vec{K}_0}(\vec{r}) \equiv \mathrm{e}^{\mathrm{i}(\vec{k}+\vec{K}_0).\vec{r}}u_{\vec{k}+\vec{K}_0}(\vec{r}) = \mathrm{e}^{\mathrm{i}(\vec{k}+\vec{K}_0).\vec{r}}\mathrm{e}^{-\mathrm{i}\vec{K}_0.\vec{r}}u_{\vec{k}}(\vec{r}) = \psi_{\vec{k}}(\vec{r}) \ , \tag{30.51}$$

confirmant que deux fonctions de Bloch dont les vecteurs d'onde diffèrent d'un vecteur de $\tilde{\mathcal{B}}$ sont en fait une seule et même fonction : on épuise bien l'ensemble de tous les états physiques distincts en confinant $\vec{k}$ à une maille élémentaire du réseau réciproque. Une fonction de Bloch est *périodique* par rapport à son paramètre $\vec{k}$, une propriété qui se reporte évidemment sur la valeur propre associée $\varepsilon(\vec{k})$:

$$\boxed{\psi_{\vec{k}+\vec{K}}(\vec{r}) = \psi_{\vec{k}}(\vec{r}) \qquad \varepsilon(\vec{k}+\vec{K}) = \varepsilon(\vec{k}) \qquad \forall\vec{K}\in\tilde{\mathcal{B}}} \tag{30.52}$$

Quand on adopte les conditions cycliques de Born - von Kármán, les valeurs de $\vec{k}$ sont de la forme (30.31). Par contruction, le cristal est de taille finie et contient $N = N_1N_2N_3$ mailles élémentaires, entraînant qu'il existe autant de valeurs distinctes possibles pour $\vec{k}$, pas plus. De fait, si m_j est plus grand que N_j, il existe deux entiers p_j et r_j tels que :

$$m_j = p_jN_j + r_j \qquad (r_j < N_j) \ ; \tag{30.53}$$

[23] à l'exception de $\vec{K}=0$ en vertu de (30.39).

À la réflexion, il n'est pas surprenant que la somme sur $\vec{q}$ ($=\vec{k}$) donnant le développement de Fourier de $\psi_{\vec{k}}$ ne contienne que des vecteurs d'onde équivalents par translation de $\tilde{\mathcal{B}}$: $\vec{k}$ est un bon nombre quantique et H, ou l'un de ses avatars, ne peut avoir d'éléments de matrice entre deux états associés à deux $\vec{k}$ inéquivalents, donnant deux valeurs propres distinctes pour un même (et quelconque) opérateur de translation.

reportant dans (30.31), $\vec{k}$ prend la forme :

$$\vec{k} = \sum_{j=1}^{3} \frac{p_j N_j + r_j}{N_j} \vec{b}_j = \sum_{j=1}^{3} (p_j + \frac{r_j}{N_j}) \vec{b}_j \equiv \vec{K} + \vec{k}_1 \ , \tag{30.54}$$

$\vec{K}$ désignant toujours un vecteur du réseau réciproque. Comme on vient de le voir, (30.51), deux fonctions $\psi_{\vec{k}}$ et $\psi_{\vec{k}+\vec{K}}$ représentent un seul et même état physique : le vecteur d'onde $\vec{k}$ paramétrant les fonctions de Bloch et les valeurs propres $\varepsilon(\vec{k}\,)$ peut être toujours restreint à une maille primitive du réseau réciproque. L'usage est de choisir la maille de Wigner - Seitz de $\tilde{\mathcal{B}}$, traditionnellement alors appelée *première zone de Brillouin*, en abrégé BZ1.

30.3 Premières conséquences du théorème de Bloch

30.3.1 Équation pour la fonction $u_{\vec{k}}$

Toute fonction propre d'un Hamiltonien périodique dans l'espace peut donc être cherchée sous la forme d'une fonction de Bloch :

$$\psi_{\vec{k}}(\vec{r}\,) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\, u_{\vec{k}}(\vec{r}\,) \ , \tag{30.55}$$

avec :

$$u_{\vec{k}}(\vec{r} + \vec{R}\,) = u_{\vec{k}}(\vec{r}\,) \ . \tag{30.56}$$

Tout revient donc à trouver la fonction $u_{\vec{k}}(\vec{r}\,)$; il est facile d'écrire l'équation qu'elle doit satisfaire. Partant de (30.42) et injectant la forme (30.55), on trouve :

$$-\frac{\hbar^2}{2m}\Big[(\mathrm{i}\vec{k}\,)^2 u_{\vec{k}} + 2\mathrm{i}\vec{k}.\vec{\nabla} u_{\vec{k}} + \vec{\nabla}^2 u_{\vec{k}}\Big] + V(\vec{r}\,)u_{\vec{k}}(\vec{r}\,) = \varepsilon(\vec{k}\,)\, u_{\vec{k}}(\vec{r}\,) \ , \tag{30.57}$$

soit, puisque $-\mathrm{i}\vec{\nabla} = \hbar^{-1}\vec{p}$:

$$\boxed{\frac{1}{2m}(\vec{p} + \hbar\vec{k}\,)^2\, u_{\vec{k}}(\vec{r}) + V(\vec{r}\,)u_{\vec{k}}(\vec{r}\,) = \varepsilon(\vec{k}\,)u_{\vec{k}}(\vec{r}\,)} \tag{30.58}$$

En vertu de la périodicité de $u_{\vec{k}}$, (30.56), il suffit de résoudre (30.58) à l'intérieur d'une maille primitive de $\mathcal{B}$, et de prendre en compte des conditions aux limites sur la frontière du domaine de résolution. On sait que toute fonction propre doit être continue et à dérivée continue tant que le potentiel ne présente pas de saut d'amplitude infinie, ce que l'on suppose ; les conditions aux limites sont donc ici :

$$\boxed{u_{\vec{k}} \text{ et } \vec{\nabla} u_{\vec{k}} \text{ continues sur la frontière d'une maille primitive}} \tag{30.59}$$

Comme toujours, l'application de conditions aux limites produit la quantification des valeurs propres ε, désormais affublées de l'indice discret n, tout comme les fonctions propres. Une notation précise est donc :

$$H_u(\vec{r}, \vec{p}; \vec{k}\,)\, u_{n\vec{k}}(\vec{r}\,) = \varepsilon_n(\vec{k}\,)\, u_{n\vec{k}}(\vec{r}\,) \ , \tag{30.60}$$

avec, d'après (30.58) :

$$\boxed{H_u(\vec{r}, \vec{p}; \vec{k}\,) \overset{\text{déf}}{=} \frac{1}{2m}\,(\vec{p} + \hbar\vec{k}\,)^2 + V(\vec{r}\,)} \tag{30.61}$$

Les valeurs propres $\varepsilon_n(\vec{k}\,)$ dépendent évidemment de $\vec{k}$; pour chaque valeur de n, les valeurs de $\vec{k}$ formant un ensemble presque dense, $\varepsilon_n(\vec{k}\,)$ varie quasi-continûment et définit une *bande d'énergie* s'étendant de $\inf_{\vec{k}} \varepsilon_n$ à $\sup_{\vec{k}} \varepsilon_n$.

La caractéristique essentielle de cette bande est sa loi de dispersion, définie par la variation de ε_n avec $\vec{k}$. L'existence des bandes d'énergie est clairement une *conséquence* directe de la symétrie de translation. Noter un fait majeur : la fonction $\varepsilon_n(\vec{k})$ est ce qu'elle est et, *a priori*, peut être fort différente de la loi de dispersion libre $\varepsilon^{(0)}(\vec{k})$ définie en (30.46). On verra que, même quand le potentiel de réseau $V(\vec{r})$ est petit, la fonction $\varepsilon_n(\vec{k})$ possède des propriétés qui la démarquent *qualitativement* de $\varepsilon^{(0)}(\vec{k})$ (en un sens précisé ultérieurement, elle a des *trous*). Chacun de ces modes propres possède évidemment la périodicité en $\vec{k}$ dans toute translation du réseau réciproque :

$$\boxed{\forall n, \forall \vec{K} \in \tilde{\mathcal{B}} \ : \ \psi_{n\vec{k}+\vec{K}}(\vec{r}\,) = \psi_{n\vec{k}}(\vec{r}\,) \qquad \varepsilon_n(\vec{k} + \vec{K}) = \varepsilon_n(\vec{k}\,)} \tag{30.62}$$

conformément à (30.52).

30.3.2 Impulsion de l'électron dans le cristal

Le vecteur $\hbar\vec{k}$ apparaissant dans le théorème de Bloch ressemble indéniablement à une impulsion au sens usuel. Il ne s'agit cependant pas de la vraie impulsion de l'électron, quoique la terminologie traditionnelle l'appelle *impulsion de l'électron dans le cristal*, ou *pseudo-impulsion*. En effet, la véritable impulsion, notée $\vec{p}$ comme d'habitude, est représentée par l'opérateur $-\mathrm{i}\hbar\vec{\nabla}$, qui n'est pas une constante du mouvement puisque l'électron est soumis à la force $-\vec{\nabla}V$ (techniquement : $[H, \vec{p}] \neq 0$). Au contraire, $\vec{k}$ est un *bon nombre quantique*, puisqu'il étiquette proprement les états propres d'un électron. D'ailleurs, $\psi_{n\vec{k}}$ n'est pas propre de $\vec{p}$:

$$\vec{p}\psi_{n\vec{k}} = -\mathrm{i}\hbar\vec{\nabla}\big(\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\, u_{n\vec{k}}(\vec{r}\,)\big) = -\mathrm{i}\hbar\big[\mathrm{i}\vec{k}\, u_{n\vec{k}} + \big(\vec{\nabla}u_{n\,\vec{k}}\big)\big]\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} = \hbar\vec{k}\psi_{n\vec{k}} - \mathrm{i}\hbar\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}\big(\vec{\nabla}\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\psi_{n\vec{k}}\big) \ . \tag{30.63}$$

Le second terme n'a aucune raison d'être proportionnel à $\psi_{n\vec{k}}$ – il ne disparaît que si la fonction $u_{n\vec{k}}(\vec{r}\,)$ est constante, c'est-à-dire pour un électron *libre*. Enfin, comme le vecteur $\vec{k}$ est défini à un vecteur $\vec{K}$ près (tous les modes propres ont la symétrie de translation décrite par (30.62)), on peut ajouter à $\vec{k}$ n'importe quel $\vec{K} \in \tilde{\mathcal{B}}$ sans changer l'état physique ; à lui seul, cet arbitraire atteste du fait que $\hbar\vec{k}$ ne peut être une impulsion physique au sens usuel.

30.3.3 Vitesse d'un électron de Bloch

On appelle désormais *électron de Bloch* un électron décrit par une fonction d'onde de Bloch, soit $\psi_{n\vec{k}}(\vec{r}) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} u_{n\vec{k}}(\vec{r})$. Il existe une relation remarquable, et très simple, entre la vitesse moyenne d'un tel électron et l'énergie propre associée à $\psi_{n\vec{k}}$:

$$\boxed{\langle \vec{v}\,\rangle_{n\vec{k}} \stackrel{\text{déf}}{=} \langle \psi_{n\vec{k}}|\frac{\vec{p}}{m}|\psi_{n\vec{k}}\rangle = \hbar^{-1}\vec{\nabla}_{\vec{k}}\,\varepsilon_n \equiv \vec{\nabla}_{\hbar\vec{k}}\,\varepsilon_n} \tag{30.64}$$

Pour établir ce résultat, on utilise le théorème de Hellmann - Feynman, déjà rencontré au chapitre 28, éq. (28.143), à propos du théorème du Viriel. Ici, le Hamiltonien pour la fonction $u_{\vec{k}}$ est H_u donné par l'expression (30.61), et dépend paramétriquement du vecteur $\vec{k}$; le théorème de Hellmann - Feynman s'écrit ici :

$$\vec{\nabla}_{\vec{k}}\varepsilon_n(\vec{k}\,) = \langle u_{n\vec{k}}|\vec{\nabla}_{\vec{k}}H_u|u_{n\vec{k}}\rangle = \langle u_{n\vec{k}}|\frac{\hbar}{m}(\vec{p}+\hbar\vec{k}\,)\,|u_{n\vec{k}}\rangle = \frac{\hbar^2}{m}\langle u_{n\vec{k}}|-\mathrm{i}\vec{\nabla}_{\vec{r}}+\vec{k}|u_{n\vec{k}}\rangle\ . \tag{30.65}$$

On a $(-\mathrm{i}\vec{\nabla}_{\vec{r}} + \vec{k}\,)u_{n\vec{k}} \equiv (-\mathrm{i}\vec{\nabla}_{\vec{r}} + \vec{k}\,)\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\psi_{n\vec{k}}(\vec{r}) = -\mathrm{i}\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\vec{\nabla}_{\vec{r}}\psi_{n\vec{k}}$, d'où :

$$\vec{\nabla}_{\vec{k}}\,\varepsilon_n(\vec{k}\,) = -\mathrm{i}\frac{\hbar^2}{m}\langle u_{n\vec{k}}|\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}\,\vec{\nabla}_{\vec{r}}\,|\psi_{n\vec{k}}\rangle \equiv \frac{\hbar}{m}\,\langle \psi_{n\vec{k}}|\vec{p}\,|\psi_{n\vec{k}}\rangle\ , \tag{30.66}$$

soit la relation annoncée :

$$\boxed{\langle \vec{v}\,\rangle_{n\vec{k}} \stackrel{\text{déf}}{=} \langle \psi_{n\vec{k}}|\frac{\vec{p}}{m}\,|\psi_{n\vec{k}}\rangle = \frac{1}{\hbar}\,\vec{\nabla}_{\vec{k}}\,\varepsilon_n(\vec{k}\,)} \tag{30.67}$$

On remarque au passage l'analogie avec une équation de Hamilton :

$$\langle \vec{v}\,\rangle = \frac{\partial \varepsilon_n}{\partial(\hbar\vec{k}\,)} \qquad \longleftrightarrow \qquad \dot{q} = \frac{\partial H}{\partial p}\ , \tag{30.68}$$

la pseudo-impulsion $\hbar\vec{k}$ jouant le rôle d'un moment conjugué et $\varepsilon_n(\vec{k}\,)$ celui d'un Hamiltonien. Par ailleurs, l'égalité (30.67), trivialement vérifiée pour un électron libre, peut être prise comme *définition* de la vitesse moyenne d'un électron dans toute théorie à une particule.

Dans tout état propre – qui produit un état stationnaire par simple adjonction du facteur de phase $\mathrm{e}^{-\mathrm{i}\hbar^{-1}\varepsilon_n(\vec{k}\,)t}$ –, la vitesse d'un électron de Bloch est constante dans le temps. L'hypothèse d'un réseau d'ions fixes est évidemment essentielle pour ce point : si l'on introduit les vibrations du réseau, les états de Bloch ne sont plus stationnaires et la vitesse moyenne d'un électron peut de fait varier – et d'ailleurs se dégrader – par collisions (création - annihilation de phonons).

30.3.4 Symétrie de la fonction $\varepsilon(\vec{k})$

Pour simplifier, on omet ci-dessous l'indice n et on considère exclusivement une bande d'énergie $\varepsilon(\vec{k})$. Il s'agit de montrer que, en l'absence de champ magnétique, on a :

$$\boxed{\varepsilon(-\vec{k}) = \varepsilon(\vec{k}) \ , \qquad \vec{\nabla}\varepsilon(\frac{\vec{K}}{2}) = 0 \qquad \forall \vec{K} \in \tilde{\mathcal{B}}} \tag{30.69}$$

avec en plus l'hypothèse que la fonction $\varepsilon(\vec{k})$ est partout deux fois dérivable.

La fonction de Bloch $\psi_{\vec{k}}$ satisfait $H\psi_{\vec{k}} = \varepsilon(\vec{k})\psi_{\vec{k}}$. En l'absence du spin, l'opération de renversement du temps se traduit par la conjugaison complexe de la fonction d'onde (voir chapitre 17, éq. (17.176)) ; par ailleurs, lorsque le Hamiltonien est invariant par renversement du temps, une fonction propre ψ et sa transformée ψ^* correspondent à la même valeur propre E. Dans le contexte actuel, ces résultats généraux s'expriment comme suit.

H étant supposé invariant par renversement du temps[24], une fonction de Bloch $\psi_{\vec{k}}(\vec{r})$ et sa complexe conjuguée $\psi^*_{\vec{k}}(\vec{r})$ ont la même énergie $\varepsilon(\vec{k})$. Par ailleurs :

$$\psi_{\vec{k}}(\vec{r}) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}u_{\vec{k}}(\vec{r}) \ , \qquad \psi^*_{\vec{k}}(\vec{r}) = \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}}u^*_{\vec{k}}(\vec{r}) \ . \tag{30.70}$$

Étant une fonction de Bloch nantie du préfacteur $\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}} \equiv \mathrm{e}^{\mathrm{i}(-\vec{k}).\vec{r}}$, $\psi^*_{\vec{k}}(\vec{r})$ est du genre $u_{-\vec{k}}(\vec{r})\mathrm{e}^{\mathrm{i}(-\vec{k}).\vec{r}}$. La comparaison avec (30.70) permet d'identifier $u_{-\vec{k}}(\vec{r})$ et $u^*_{\vec{k}}(\vec{r})$, d'où :

$$\boxed{\psi^*_{\vec{k}}(\vec{r}) = \psi_{-\vec{k}}(\vec{r})} \tag{30.71}$$

Ces deux fonctions, transformées l'une de l'autre par renversement du temps, ont la même énergie ; mais $\psi_{-\vec{k}}(\vec{r})$ satisfait $H\psi_{-\vec{k}} = \varepsilon(-\vec{k})\psi_{-\vec{k}}$, d'où l'égalité :

$$\boxed{\varepsilon(-\vec{k}) = \varepsilon(\vec{k})} \tag{30.72}$$

$\varepsilon_{\vec{k}}$ étant une fonction paire, sa dérivée est une fonction impaire :

$$\vec{\nabla}\varepsilon(-\vec{k}) = -\vec{\nabla}\varepsilon(\vec{k}) \ . \tag{30.73}$$

En outre, $\varepsilon(\vec{k})$ est une fonction périodique en $\vec{K}$, tout comme son gradient :

$$\varepsilon(\vec{k} - \vec{K}) = \varepsilon(\vec{k}) \qquad \text{et} \qquad \vec{\nabla}\varepsilon(\vec{k} - \vec{K}) = \vec{\nabla}\varepsilon(\vec{k}) \ ; \tag{30.74}$$

[24] C'est toujours le cas quand $H = \frac{\vec{p}^2}{2m} + V(\vec{r})$, une expression qui exclut la présence d'un champ magnétique. L'invariance de H par renversement du temps vient du fait que H est *quadratique* par rapport à la vitesse. Au contraire, un champ magnétique implique forcément une boucle de courant, c'est-à-dire un moment cinétique, de spin ou orbital, qui s'inverse dans le changement $t \to -t$.

pour les vecteurs d'onde tels que $\vec{k} - \vec{K} = -\vec{k}$, soit $\vec{k} = \frac{\vec{K}}{2}$, les égalités (30.73) et (30.74) donnent simultanément[25] :

$$\vec{\nabla}\varepsilon\Big(-\frac{\vec{K}}{2}\Big) = -\vec{\nabla}\varepsilon\Big(\frac{\vec{K}}{2}\Big) \quad \text{et} \quad \vec{\nabla}\varepsilon\Big(-\frac{\vec{K}}{2}\Big) = \vec{\nabla}\varepsilon\Big(\frac{\vec{K}}{2}\Big) \quad \Longleftrightarrow \quad \vec{\nabla}\varepsilon\Big(\frac{\vec{K}}{2}\Big) = 0 \ . \tag{30.75}$$

$\varepsilon(\vec{k}\,)$ est *extremum* en ces points, en particulier en $\vec{k} = 0$.

30.3.5 Surface de Fermi

La surface de Fermi est définie et existe dans tout modèle à électrons indépendants (au sens large) – son observation expérimentale dans un métal est d'ailleurs usuellement considérée comme un test certifiant la pertinence d'une telle approximation –, quel que soit le potentiel à un corps ressenti par un électron dans le modèle effectif considéré. Dans ces conditions, la construction de l'état fondamental d'un ensemble de N électrons s'effectue facilement, en conformité avec le Principe de Pauli : on prend les états électroniques un par un, par ordre d'énergie croissante, et on y place deux électrons de spins opposés, jusqu'à épuisement du stock d'électrons. L'énergie maximum ainsi atteinte s'appelle énergie de Fermi, $\varepsilon_{\rm F}$. Si les niveaux sont repérés par un vecteur d'onde $\vec{k}$, et s'il n'existe pas de recouvrement entre les bandes[26], les niveaux occupés de plus haute énergie sont tels que :

$$\varepsilon_{>}(\vec{k}\,) = \varepsilon_{\rm F} \ , \tag{30.76}$$

où $\varepsilon_{>}(\vec{k}\,)$ est la loi de dispersion de la seule et unique plus haute bande participant à l'étape finale du remplissage. Cette égalité définit une surface dans l'espace $\vec{k}$, appelée *surface de Fermi.*

La situation la plus simple est celle des électrons libres, la loi de dispersion étant la fonction $\varepsilon^{(0)}(\vec{k}\,)$ définie en (30.46) de sorte que la surface de Fermi a pour équation :

$$\frac{\hbar^2\vec{k}^2}{2m} = \varepsilon_{\rm F} \ . \tag{30.77}$$

C'est donc une sphère, ayant un certain rayon $k_{\rm F}$; la valeur de $k_{\rm F}$ a déjà été trouvée lors de l'étude de l'énergie de cohésion d'un métal (voir chapitre 29, éq. (29.27)), et vaut (V est ici le volume du cristal) :

$$\boxed{k_{\rm F} = \Big(3\pi^2\,\frac{N}{V}\Big)^{1/3} \equiv (3\pi^2\rho)^{1/3} \sim \frac{1}{d}} \tag{30.78}$$

où d est l'ordre de grandeur de la distance moyenne entre deux électrons.

[25]La discussion en cours exclut les cas où $\varepsilon(\vec{k}\,)$ présente des points anguleux puisque la fonction $\varepsilon(\vec{k}\,)$ a été supposée deux fois dérivable partout.

[26]Quand il y a recouvrement en énergie entre les bandes, il peut y avoir plusieurs bandes incomplètement remplies, chacune jusqu'à la même énergie $\varepsilon_{\rm F}$ (penser aux vases communicants).

Pour un gaz d'électrons soumis à un potentiel périodique, la surface de Fermi est définie par considération des plus hautes bandes occupées. Toujours en l'absence de recouvrement, la surface de Fermi est définie par l'équation (30.76), mais $\varepsilon(\vec{k}\,)$ est maintenant la loi de dispersion relative à cette bande. Toutefois, comme il ne faut compter qu'une fois chaque état et que deux états ne différant que par un vecteur $\vec{K}$ de $\tilde{\mathcal{B}}$ sont identiques, $\vec{k}$ doit être restreint à une maille primitive de $\tilde{\mathcal{B}}$, la première zone de Brillouin par exemple. Il est donc utile de préciser :

$$\varepsilon(\vec{k}\,) = \varepsilon_{\mathrm{F}} \qquad \forall \vec{k} \in \mathrm{BZ1} \ . \tag{30.79}$$

Cela étant, et comme $\varepsilon(\vec{k}\,)$ est périodique dans $\tilde{\mathcal{B}}$, il est parfois utile de définir la surface de Fermi complète, obtenue en appliquant à la surface définie par (30.79) toutes les translations du réseau réciproque.

La valeur de ε_{F} est déterminée en intégrant la densité d'états $\mathcal{D}(E)$ (définie plus loin) jusqu'à ε_{F} et en écrivant que cette intégrale est égale au nombre d'électrons à caser. Précisément, pour un cristal à D dimensions de volume V, on écrit :

$$2\frac{V}{(2\pi)^D}\int_0^{\varepsilon_{\mathrm{F}}} \mathrm{d}E \int_{\mathbb{R}^D} \mathrm{d}^D k \, \delta(\varepsilon(\vec{k}\,) - E) = N_{\mathrm{e}} \ , \tag{30.80}$$

où N_{e} est le nombre d'électrons à répartir dans la bande ayant la dispersion $\varepsilon(\vec{k}\,)$; de façon équivalente, ρ_{e} étant la densité électronique :

$$\int_0^{\varepsilon_{\mathrm{F}}} \mathrm{d}E \int_{\mathbb{R}^D} \mathrm{d}^D k \, \delta(\varepsilon(\vec{k}) - E) = 2^{D-1}\pi^D \rho_{\mathrm{e}} \ . \tag{30.81}$$

Compte tenu de la périodicité de $\varepsilon(\vec{k}\,)$ dans $\tilde{\mathcal{B}}$, la surface de Fermi complète s'obtient par les translations de $\tilde{\mathcal{B}}$ de l'objet défini par (30.79) ; cette construction est utile pour l'analyse des problèmes de transport et permet, notamment, de distinguer le caractère ouvert ou fermé des orbites (trajectoires) semi-classiques. La figure 30.1 présente deux "surfaces" de Fermi dans le cas d'un réseau carré bidimensionnel pour deux valeurs très différentes du moment de Fermi k_{F}. Le grand carré est la première zone de Brillouin, de côté $\frac{2\pi}{a}$; Γ est le centre de la zone, ses côtés sont des "plans" de Bragg. Lorsque $k_{\mathrm{F}} \ll \frac{2\pi}{a}$ ($\equiv k_{\mathrm{F1}}$), la surface de Fermi est très proche de la "sphère" (cercle) libre (cercle central sur la figure). Au contraire, si la densité d'électrons est assez élevée, $k_{\mathrm{F}} \gtrsim \frac{2\pi}{a}$ ($\equiv k_{\mathrm{F2}}$), la sphère libre (cercle en pointillés) se déforme considérablement près des plans de Bragg. Par des translations horizontales et verticales de $\pm\frac{2\pi}{a}$, les arcs déformés de la sphère libre peuvent être transportés à l'intérieur de BZ1, d'où les poches dans les coins et les arcs semi-elliptiques au milieu de chaque côté[27]. La surface de Fermi elle-même est la réunion de tous les arcs en gras, tous périodisés dans l'espace réciproque par une translation quelconque d'un vecteur de $\mathcal{B}$.

Il existe une différence topologique importante entre ces deux cas : quand la densité est faible, k_{F1}, la surface de Fermi périodisée est constituée de quasi-sphères disjointes, chacune étant entièrement contenue dans BZ1 ; les électrons situés dans l'une

[27] Au total, la surface du cercle en pointillés est égale à la surface du domaine situé dans BZ1 et délimité par les arcs définissant les poches et les demi-ellipses.

des sphères y sont prisonniers et ne peuvent sortir de la zone correspondante : toute trajectoire d'énergie continue est forcément fermée. Au contraire, quand la densité est suffisamment grande, $k_{\mathrm{F}2}$, la surface de Fermi périodisée contient des canaux de communication d'une zone à l'autre, permettant le *déconfinement* d'un électron initialement situé dans l'une quelconque d'entre elles.

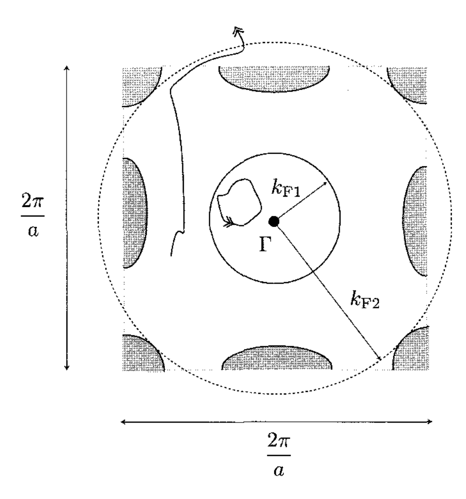


Figure 30.1: Surface de Fermi pour un réseau carré bidimensionnel pour deux valeurs du moment de Fermi k_{F}. Les deux lignes symbolisent deux trajectoires, l'une forcément fermée, l'autre ouverte (voir le texte).

30.4 Applications choisies du théorème de Bloch

Le théorème de Bloch énonce un résultat exact sur la nature des modes propres d'une particule dans un potentiel périodique[28], et montre notamment que les valeurs propres $\varepsilon_n(\vec{k}\,)$ sont des fonctions (périodiques dans $\tilde{\mathcal{B}}$) du vecteur $\vec{k}$: ce sont les lois de dispersion de la particule dans le cristal, donnant la variation de son énergie en fonction de l'impulsion, plus précisément de ce qui tient lieu d'impulsion – voir les mises en garde terminologiques du chapitre précédent. Il s'agit maintenant de présenter deux applications importantes de ce théorème, qui définissent deux cadres d'approximation adaptés à deux situations physiques extrêmes, identifiées par les caractères de ces lois de dispersion :

1. l'approximation des électrons presque libres, correspondant au cas où, dans le solide, les électrons sont seulement faiblement perturbés par le potentiel de réseau par comparaison à leur état libre. Il en résulte que la loi de dispersion ressemble à la parabole libre définie en (30.46) ; elle est donc fortement dispersée et on verra qu'elle est constituée d'arcs séparés par des petits *gaps*. Ceci correspond à des bandes

[28] Pour une belle discussion détaillée d'un potentiel périodique à une dimension, voir [4], chapitre 6.

larges et à une forte délocalisation spatiale des électrons (onde plane faiblement modulée). Cette situation prévaut dans les métaux et les semi-conducteurs.

2. À l'opposé, l'approximation des liaisons fortes décrit le cas où chaque électron ressent, à l'ordre zéro, le potentiel localisé d'un atome (ou ion) du réseau. À cet ordre, le spectre d'énergie est celui d'un atome : des niveaux discrets, bien séparés. En fait, chaque électron peut se déplacer d'un atome à l'autre, ce qui donne une petite dispersion aux niveaux issus des niveaux atomiques. Il en résulte que les bandes d'énergie sont étroites et séparées par des grands *gaps*. Lorsque la plus haute bande est pleine, cette situation est typique des isolants.

Ces deux situations sont apparues dans le problème 17.8.6, dans le cas très particulier du modèle de Kronig - Penney.

30.4.1 L'approximation des électrons presque libres

On adopte donc ici l'hypothèse suivant laquelle le potentiel de réseau $V(\vec{r})$ constitue une petite perturbation sur les états propres libres :

$$\psi_{\vec{k}}^{(0)}(\vec{r}) = C\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \ , \qquad \varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2\vec{k}^2}{2m} \ , \tag{30.82}$$

ce qui autorise d'emblée à envisager un traitement de type perturbatif, avec les précautions et les hypothèses usuelles.

Ces états propres à l'ordre zéro sont *qualitativement* différents des vrais états propres de l'électron dans le cristal, puisque ceux-ci doivent avoir la symétrie exigée par le théorème de Bloch et ses conséquences en raison de la symétrie de translation *discrète* – ce n'est pas le cas des états libres (30.82), qui sont caractéristiques d'une symétrie de translation continue, celle d'une particule libre. De surcroît, dans (30.82), $\hbar\vec{k}$ est visiblement la *vraie* impulsion de l'électron.

Pour tout dire, on rencontre à nouveau un bel exemple du fait que la symétrie fonctionne par tout ou rien et que des considérations *quantitatives* – du genre de celles que l'on énonce pour mettre en évidence un petit paramètre par exemple – appartiennent à un autre plan. Que le potentiel de réseau soit faible ou intense ne change rien au fait que la symétrie de translation devient discrète même si $V(\vec{r})$ est infiniment petit[29]. Il s'agit bien d'une *brisure* de symétrie, puisque la symétrie est devenue discrète sous l'effet du réseau de pas fini, aussi petit que soit $V(\vec{r})$. Le groupe des translations discrètes est

[29]On a toujours intérêt à faire jouer la symétrie le plus tôt possible. Notamment, l'introduction explicite de vecteurs ayant la symétrie de la perturbation produit des matrices diagonales par blocs. Il arrive même, quand la brisure de symétrie est assez forte, qu'à l'intéreur d'un sous-espace dégénéré, la projection de l'opérateur de perturbation dans ce sous-espace est représentée par une matrice complètement diagonale ; la simple lecture de ces éléments diagonaux donne alors immédiatement, sans aucun calcul supplémentaire, les corrections d'énergie au premier ordre.

visiblement un sous-groupe des translations continues ; on retrouve indiscutablement ce dernier en faisant tendre vers zéro le pas du réseau.

Comme on l'a vu précédemment, l'existence du potentiel de réseau entraîne que les fonctions propres et les valeurs propres sont périodiques dans $\tilde{\mathcal{B}}$:

$$\boxed{\psi_{n\vec{k}-\vec{K}}(\vec{r}) = \psi_{n\vec{k}}(\vec{r}) \ , \qquad \varepsilon_n(\vec{k}-\vec{K}) = \varepsilon_n(\vec{k}) \qquad \forall \vec{K} \in \tilde{\mathcal{B}}} \tag{30.83}$$

Ceci signifie qu'il suffit de les connaître à l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$ et, dans la suite, on pourra toujours, si besoin est, choisir le vecteur $\vec{k}$ dans la première zone de Brillouin[30]. Compte tenu de l'hypothèse sur la petitesse de V, on va utiliser une méthode de perturbation ; il convient donc de préciser complètement le problème, tel qu'il se pose à l'ordre zéro, le reformulant dans la perspective du traitement choisi. Par ailleurs, on sait que la technique de perturbation, avant d'être mise en œuvre, exige l'analyse préalable de la dégénérescence du problème non-perturbé.

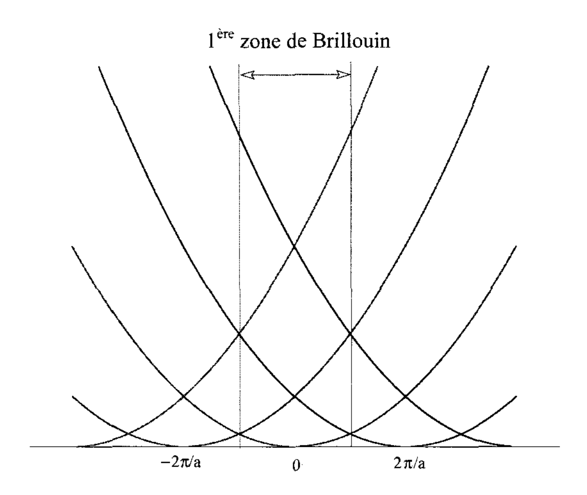


Figure 30.2: Schéma en zones réduites pour un réseau $\mathcal{B}$ unidimensionnel de pas a. Les différentes paraboles représentent les $\varepsilon^{(0)}_{\vec{K}}(\vec{k})$ (voir (30.85)). La première zone de Brillouin s'étend entre $\pm\frac{\pi}{a}$.

Les solutions exactes sont périodiques dans $\tilde{\mathcal{B}}$, ainsi que l'exprime (30.83). D'un autre côté, comme V reste en toute hypothèse petit, on s'attend nécessairement à ce que l'on ait toujours :

$$\varepsilon(\vec{k}) \approx \varepsilon^{(0)}(\vec{k}) \ . \tag{30.84}$$

La bonne façon de formuler le problème à l'ordre zéro est donc de prolonger par périodisation, conformément à (30.83), tous les arcs de la parabole libre. Ceci se fait en translatant tous les arcs de la parabole libre qui sont en-dehors de cette zone, obtenant alors

[30] Comme on l'a vu, dès qu'il existe un potentiel périodique, $\vec{k}$ n'est défini physiquement que *modulo* $\vec{K} \in \tilde{\mathcal{B}}$, tout comme un angle est défini à 2π près ; on peut dire que l'équivalent de BZ1 est le segment $[-\pi, +\pi]$ pour une fonction 2π-périodique.

le schéma dit en zones réduites (voir figure 30.2). Dans cette représentation, l'énergie d'un électron en fonction de $\vec{k}$ est une fonction infiniment multivaluée[31]. Avec la même notation que précédemment, $\varepsilon^{(0)}(\vec{k})$, le prolongement périodique revient donc à définir une infinité de lois paraboliques $\varepsilon^{(0)}_{\vec{K}}$ telles que :

$$\varepsilon^{(0)}_{\vec{K}}(\vec{k}) \stackrel{\text{déf}}{=} \varepsilon^{(0)}(\vec{k} - \vec{K}) \equiv \frac{\hbar^2}{2m}(\vec{k} - \vec{K})^2 \qquad \forall \vec{K} \in \tilde{\mathcal{B}} \ . \tag{30.85}$$

Ceci constitue la *vraie* solution exacte dans la limite $V(\vec{r}) = 0+$, et permet l'analyse préalable de la dégénérescence en vue d'un traitement perturbatif.

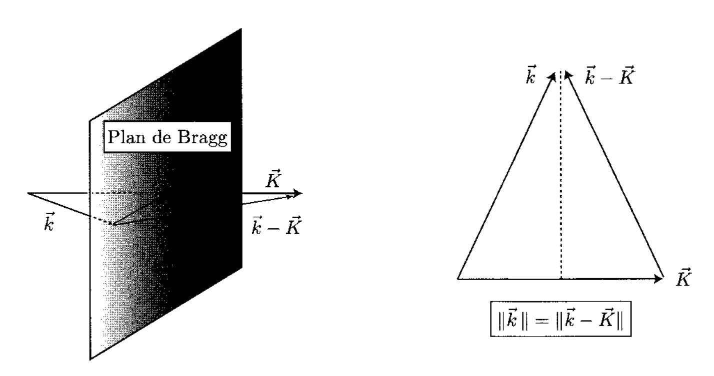


Figure 30.3: Schéma de définition du plan de Bragg d'un vecteur $\vec{K} \in \tilde{\mathcal{B}}$.

Il y aura dégénérescence stricte à chaque fois que toutes ces paraboles se croisent, c'est-à-dire lorsque :

$$\varepsilon^{(0)}_{\vec{K}_1}(\vec{k}) = \varepsilon^{(0)}_{\vec{K}_2}(\vec{k}) \quad \Longleftrightarrow \quad \varepsilon^{(0)}(\vec{k} - \vec{K}) = \varepsilon^{(0)}(\vec{k}) \ , \tag{30.86}$$

où $\vec{K} = \vec{K}_2 - \vec{K}_1$ est encore un vecteur de $\tilde{\mathcal{B}}$. Comme $\varepsilon^{(0)}(\vec{k})$ ne dépend que du module des vecteurs d'onde, cette égalité (30.86) est satisfaite à chaque fois que[32] :

$$\boxed{\|\vec{k}\| = \|\vec{K} - \vec{k}\|} \tag{30.87}$$

On peut d'ailleurs dire les choses autrement ; pour un électron libre, la condition de dégénérescence est :

$$\varepsilon^{(0)}(\vec{k}) = \varepsilon^{(0)}(\vec{k}\,') \ . \tag{30.88}$$

[31]On peut se demander ce que devient la périodisation de $\varepsilon^{(0)}_{\vec{K}}$ quand on en revient au cas strictement libre. De fait, la limite de l'électron libre se retrouve comme suit : il suffit de faire tendre vers zéro la maille primitive du réseau, ce qui à la limite produit bien un potentiel $V(\vec{r})$ prenant la même valeur *partout*. Dans cette limite, la longueur de tout vecteur $\vec{K} \in \tilde{\mathcal{B}}$ tend vers l'infini et on récupère *in fine* l'unique parabole qui est la limite de l'arc de plus basse énergie dans le schéma en zones réduites.

[32]En l'absence de champ magnétique, $\varepsilon^{(0)}$ est aussi une fonction paire de $\vec{k}$; toutefois, il n'y a pas lieu de considérer la dégénérescence $\vec{k} \to -\vec{k}$ correspondante : $\vec{k}$ est un bon nombre quantique et les éléments de matrice $\langle\psi_{\vec{k}}|H|\psi_{-\vec{k}}\rangle$ sont nuls.

Cette égalité est vraie si $\vec{k}$ et $\vec{k}'$ ont même module, en particulier si $\vec{k} = \vec{k}'$. Mais, dès qu'un potentiel périodique est présent, l'égalité des vecteurs d'onde doit être comprise *modulo* $\vec{K}$. Dans cette optique, (30.88) est équivalente à (30.87).

L'équation (30.87) signifie que, $\vec{K}$ étant fixé, l'extrémité de $\vec{k}$ est dans le plan médiateur de $\vec{K}$ (voir fig. 30.3) ; ce plan bissecteur est par définition le *plan de Bragg* de $\vec{K}$, noté $\mathcal{P}_{\vec{K}/2}$. Ainsi, il apparaît une dégénérescence stricte apparaît lorsque l'extrémité de $\vec{k}$ appartient à un plan de Bragg[33] ; dans cette géométrie, on a :

$$\vec{k} = \frac{\vec{K}}{2} + \vec{q}_\perp \ , \qquad q_\perp \perp \vec{K} \tag{30.89}$$

Soit alors un vecteur primitif $\vec{b}_0$ définissant une certaine direction de $\tilde{\mathcal{B}}$, on peut distinguer deux cas :

- ou bien $\vec{K} = \text{entier pair} \times \vec{b}_0 \equiv 2n_0\vec{b}_0$. Dans ces conditions $\frac{\vec{K}}{2} = \vec{K}_0 \in \tilde{\mathcal{B}}$ et, après la bonne translation, on se retrouve à l'origine de $\tilde{\mathcal{B}}$; on parle alors de dégénérescence en *centre de zone* (points d'abscisses 0, $\pm\frac{2\pi}{a}$,... sur la figure 30.2).
- ou bien $\vec{K} = \text{entier impair} \times \vec{b}_0 \equiv (2n_0 + 1)\vec{b}_0$. Alors, $\frac{\vec{K}}{2} = \vec{K}_0 + \frac{1}{2}\vec{b}_0 \notin \tilde{\mathcal{B}}$ et la dégénérescence stricte se produit en un certain point de la frontière de la maille de Wigner - Seitz du réseau réciproque (BZ1) ; on parle alors de dégénérescence en *bord de zone* (points d'abscisses $\pm\frac{\pi}{a}$, $\pm\frac{3\pi}{a}$,... sur la figure 30.2).

Comme le paramètre $\vec{k}$ varie continûment[34], il ne serait pas satisfaisant d'utiliser presque partout dans la zone la théorie de perturbation relative au cas non-dégénéré, sauf en certains points isolés où on appliquerait le formalisme du cas dégénéré ; une telle procédure conduirait d'ailleurs à des difficultés techniques puisque les dénominateurs d'énergie peuvent devenir graduellement de plus en plus petits (voir aussi (30.90)). Il faut donc construire un schéma de calcul plus subtil que celui énoncé par la théorie des perturbations "académique"[35].

On sait que le critère convenable pour qu'une théorie de perturbation fonctionne bien est de rassembler – en les traitant sur un même pied d'égalité – tous les états d'ordre zéro $\psi_n^{(0)}$ dont les énergies $E_n^{(0)}$ ont des différences au plus comparables aux éléments de matrice de la perturbation entre ces états. Ici, le potentiel de réseau (la perturbation) est supposé faible, c'est-à-dire que ses éléments de matrice sont réputés petits par rapport à une énergie typique ; compte tenu des paramètres physiques disponibles, cette énergie ne peut être que[36] $\frac{\hbar^2}{ma^2}$. Au total, la dégénérescence doit s'apprécier par une condition

[33] sans préjuger d'autres coïncidences par ailleurs, résultant d'une haute symétrie ponctuelle.

[34] À une petite quantification près imposée par les conditions cycliques.

[35] C'est la même situation que pour un atome soumis à un champ magnétique $\vec{\mathcal{B}}$; les valeurs propres dépendant continûment de $\mathcal{B}$, le traitement de perturbation ordinaire doit être redéfini en présence d'un croisement de niveaux (voir par exemple problème 23.4.11).

[36] a est l'ordre de grandeur de la dimension linéaire d'une maille primitive.

du type :

$$\boxed{|\varepsilon^{(0)}_{\vec{K}}(\vec{k}\,) - \varepsilon^{(0)}_{\vec{K}'}(\vec{k}\,)| \lesssim |V|} \tag{30.90}$$

et le traitement de perturbation, qui exige par ailleurs :

$$|\mathcal{V}(\vec{K})| \ll \frac{\hbar^2}{ma^2} \ . \tag{30.91}$$

doit considérer *simultanément* tous les états d'ordre zéro satisfaisant (30.90).

V étant suffisamment faible, les arguments précédents permettent de délimiter deux situations typiques, nettement différenciées ($\vec{k}$ étant choisi dans BZ1) :

1. le vecteur $\vec{k}$ considéré est tel que toutes les branches $\varepsilon^{(0)}_{\vec{K}}(\vec{k}\,)$ sont bien isolées les unes des autres. Alors, une théorie de perturbation non-dégénérée est acceptable. Sur la figure 30.2, ce sont donc les points plus ou moins à mi-chemin des valeurs 0, $\pm\frac{\pi}{a}$, $\pm\frac{2\pi}{a}$, etc.

2. Au contraire, $\vec{k}$ est tel que plusieurs branches $\varepsilon^{(0)}_{\vec{K}_i}(\vec{k}\,)$ se trouvent au voisinage les unes des autres, l'étendue du voisinage se mesurant par l'ordre de grandeur du potentiel périodique conformément à (30.90). Comme on vient de le voir, ceci se produit lorsque $\vec{k}$ est près du centre ou au voisinage du plan de Bragg de $\vec{K}_i - \vec{K}_j$ (par exemple sur la frontière de BZ1, quand $\vec{K}_1 = 0$, et $\vec{K}_2$ joint l'origine à un nœud premier voisin dans $\tilde{\mathcal{B}}$, par exemple). S'il y a g telles branches, le moins que l'on puisse faire est alors de considérer *simultanément* les g fonctions $\psi^{(0)}_{\vec{k}-\vec{K}_i}$, ou plus précisément le sous-espace g $\times$ g qu'elles engendrent. S'y prendre ainsi c'est finalement étendre la recette de perturbation dégénérée dans le *voisinage* du point de dégénérescence défini par l'égalité stricte $\|\vec{k} - \vec{K}_i\| = \|\vec{k} - \vec{K}_j\|$, et non pas strictement au point de dégénérescence.

   Dans le cas le plus simple, on a g $= 2$; c'est la situation lorsque l'extrémité de $\vec{k}$ est non seulement près du plan de Bragg de $\vec{K} \stackrel{\text{déf}}{=} \vec{K}_i - \vec{K}_j$ mais encore pas trop éloigné de la direction du vecteur $\vec{K}$ lui-même. Si l'écart grandit, $\vec{k}$ peut se rapprocher de la droite d'intersection de deux plans de Bragg, et alors visiblement g > 2.

Examinons maintenant comment les deux cas se présentent naturellement, en revenant au problème exact et en le représentant sur une base d'ondes planes (les fonctions à l'ordre zéro), ce qui revient à se placer dans l'espace de Fourier. D'après (30.49) et après un simple changement des notations pour un aspect plus naturel, une fonction de Bloch admet le développement :

$$\boxed{\psi_{\vec{k}}(\vec{r}\,) = \sum_{\vec{K}\in\tilde{\mathcal{B}}} c_{\vec{k}-\vec{K}}\, \mathrm{e}^{\mathrm{i}(\vec{k}-\vec{K}).\vec{r}}} \tag{30.92}$$

Noter que chaque onde plane satisfait aussi le théorème de Bloch et ses conséquences (30.83). En termes de *kets*, on peut écrire :

$$|\psi_{\vec{k}}\rangle = \sum_{\vec{K}\in\tilde{\mathcal{B}}} c_{\vec{k}-\vec{K}}\,|f_{\vec{k}-\vec{K}}\rangle \ , \tag{30.93}$$

les fonctions $\langle\vec{r}\,|f_{\vec{k}-\vec{K}}\rangle = \mathrm{e}^{\mathrm{i}(\vec{k}-\vec{K}).\vec{r}} \stackrel{\text{déf}}{=} \psi^{(0)}_{\vec{k}-\vec{K}}(\vec{r}\,)$ étant les ondes planes non-perturbées, solutions du problème à l'ordre zéro. Les $c_{\vec{k}-\vec{K}}$ sont donc les coefficients cherchés, que l'on imagine – en vue d'un traitement perturbatif – être des séries entières en V.

Pour obtenir les coefficients $c_{\vec{k}-\vec{K}}$, il suffit de reporter les développements de Fourier (30.33) et (30.92) dans l'équation satisfaite par $\psi_{\vec{k}}(\vec{r}\,)$:

$$-\frac{\hbar^2}{2m}\,\Delta\psi_{\vec{k}}(\vec{r}\,) + V(\vec{r}\,)\,\psi_{\vec{k}}(\vec{r}\,) = \varepsilon(\vec{k}\,)\,\psi_{\vec{k}}(\vec{r}\,) \tag{30.94}$$

pour retrouver le système (30.47) qui, dans les nouvelles notations, s'écrit :

$$\boxed{[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k}-\vec{K})]c_{\vec{k}-\vec{K}} = \sum_{\vec{K}'\in\tilde{\mathcal{B}}} \mathcal{V}(\vec{K}-\vec{K}')\,c_{\vec{k}-\vec{K}'} \qquad \forall\vec{K}\in\tilde{\mathcal{B}}} \tag{30.95}$$

En choisissant convenablement le zéro d'énergie potentielle – à la valeur moyenne de $V(\vec{r}\,)$ sur une maille primitive –, la composante uniforme $\mathcal{V}(\vec{K}=0)$ est nulle ; dans ces conditions, la somme dans (30.95) court explicitement sur les valeurs de $\vec{K}' \neq \vec{K}$. Par ailleurs, ce choix entraîne qu'il n'y a pas de correction d'énergie au premier ordre en l'absence de dégénérescence. En effet, en pareil cas, la correction au premier ordre est simplement la valeur moyenne de la perturbation prise sur un état d'ordre zéro, soit ici :

$$\int_{\mathbb{R}^3} \mathrm{e}^{-\mathrm{i}(\vec{k}-\vec{K}).\vec{r}}\,V(\vec{r}\,)\,\mathrm{e}^{\mathrm{i}(\vec{k}-\vec{K}).\vec{r}}\,\mathrm{d}^3r = \int_{\mathbb{R}^3} V(\vec{r}\,)\,\mathrm{d}^3r = 0 \ . \tag{30.96}$$

Jusqu'à présent, aucune approximation n'a été faite : le système (30.95) est exact, il est tout juste une reformulation de l'équation aux valeurs propres (30.94) dans l'espace de Fourier. Inexploitable en tant que tel dans le cas général, on en effectue maintenant l'analyse perturbative.

Si $V \equiv 0$, le système (30.95) se réduit à :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k}-\vec{K})]\,c_{\vec{k}-\vec{K}} = 0 \ , \qquad \forall\vec{K}\in\tilde{\mathcal{B}} \ . \tag{30.97}$$

L'équation (30.97) dit que chaque coefficient $c_{\vec{k}-\vec{K}}$ est nul si $\varepsilon(\vec{k}\,) \neq \varepsilon^{(0)}(\vec{k}-\vec{K})$; à l'inverse, c'est seulement lorsque $\varepsilon(\vec{k}\,)$ est égal à $\varepsilon^{(0)}(\vec{k}-\vec{K})$, que $c_{\vec{k}-\vec{K}}$ n'est pas nul et est, à ce stade, indéterminé ; c'est une constante non nulle faute de quoi la solution serait identiquement nulle, condition rencontrée à propos de toute équation aux fonctions propres (homogène par nature). En tant que constante à la limite $V \to 0$, c'est donc un terme d'ordre $\mathcal{O}(V^0)$.

Dans le cas 1., par définition, toutes les énergies $\varepsilon^{(0)}(\vec{k}-\vec{K})$ sont, pour le $\vec{k}$ choisi, bien isolées les unes des autres. Considérons-en une, celle ayant $\vec{K}=\vec{K}_1$ et récrivons le système (30.95) en exhibant à part la ligne $\vec{K}=\vec{K}_1$:

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k}-\vec{K}_1)]\, c_{\vec{k}-\vec{K}_1} = \sum_{\vec{K}'\in\tilde{\mathcal{B}},\, \vec{K}'\neq\vec{K}_1} \mathcal{V}(\vec{K}_1-\vec{K}')\, c_{\vec{k}-\vec{K}'} \ , \tag{30.98}$$

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k}-\vec{K})]\, c_{\vec{k}-\vec{K}} = \sum_{\vec{K}'\in\tilde{\mathcal{B}},\, \vec{K}'\neq\vec{K}} \mathcal{V}(\vec{K}-\vec{K}')\, c_{\vec{k}-\vec{K}'} \ , \quad \forall \vec{K}\in\tilde{\mathcal{B}},\, \vec{K}\neq\vec{K}_1 \ . \tag{30.99}$$

Quand $V\to 0$, un seul crochet $[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k}-\vec{K})]$ figurant aux premiers membres du système (30.95) s'annule, celui de (30.98), et la valeur propre perturbée $\varepsilon(\vec{k})$ tend vers l'unique valeur isolée $\varepsilon^{(0)}(\vec{k}-\vec{K}_1)$ considérée. Bien sûr, tous les autres premiers membres de (30.99) s'annulent aussi, mais c'est alors parce que, le crochet n'étant pas nul à la limite, le coefficient $c_{\vec{k}-\vec{K}}$ lui-même tend vers zéro $\forall\,\vec{K}\neq\vec{K}_1$. Dans ces conditions, chaque $c_{\vec{k}-\vec{K}}$, $\vec{K}\neq\vec{K}_1$ est au moins d'ordre 1 en V ; il en résulte que la somme au second membre de (30.98) – qui ne contient que les $c_{\vec{k}-\vec{K}'}$, $\vec{K}'\neq\vec{K}_1$ – est au moins du second ordre[37]. En définitive, la différence d'énergie $\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k}-\vec{K}_1)$ est du second ordre[38] et la correction résultant du petit potentiel périodique est très petite. La fonction d'onde diffère de la fonction d'onde d'ordre zéro par des termes du premier ordre : l'ordre zéro est la série de Fourier (30.49) juste réduite au seul terme $\mathrm{e}^{\mathrm{i}(\vec{k}-\vec{K}).\vec{r}}$, c'est bien une onde plane. Au total, il n'y a pas de changement qualitatif notable.

Dans le cas 2., toujours avec $\vec{k}$ fixé mais pris cette fois près du centre ou du bord de zone, il existe par hypothèse un certain nombre g > 1 d'énergies quasi-dégénérées associées à des vecteurs $\vec{K}_i$ distincts :

$$|\varepsilon^{(0)}(\vec{k}-\vec{K}_i) - \varepsilon^{(0)}(\vec{k}-\vec{K}_j)| \lesssim |V| \qquad (i,\, j = 1,\, 2,\, \dots\, \mathrm{g}) \ . \tag{30.100}$$

Mettons à part toutes les lignes où ces $\vec{K}_i$ sont au premier membre de (30.95) :

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k}-\vec{K}_i)]\, c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}'\in\tilde{\mathcal{B}},\, \vec{K}'\neq\vec{K}_i} \mathcal{V}(\vec{K}_i-\vec{K}')\, c_{\vec{k}-\vec{K}'} \ , \ \ \vec{K}_i = \vec{K}_1,\, \vec{K}_2,\, \dots,\, \vec{K}_{\mathrm{g}} \ , \tag{30.101}$$

$$[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k}-\vec{K})]\, c_{\vec{k}-\vec{K}} = \sum_{\vec{K}'\in\tilde{\mathcal{B}},\, \vec{K}'\neq\vec{K}} \mathcal{V}(\vec{K}-\vec{K}')\, c_{\vec{k}-\vec{K}'} \ , \ \ \forall\,\vec{K}\in\tilde{\mathcal{B}}\backslash\{\vec{K}_i\}_{i=1,\,2,\,\dots,\,\mathrm{g}} \ . \tag{30.102}$$

Le même argument que celui développé pour le cas 1. montre alors que les g crochets $[\varepsilon(\vec{k}) - \varepsilon^{(0)}(\vec{k}-\vec{K}_i)]$ des premiers membres de (30.101) s'annulent à la limite $V=0$, donc les g coefficients $c_{\vec{k}-\vec{K}_i}$ sont tous d'ordre $\mathcal{O}(V^0)$, tous les autres étant au moins d'ordre

[37] Comme toujours quand il s'agit de perturbation ordinaire (non-singulière), on fait l'hypothèse suivant laquelle tous les développements sont en puissances entières de la perturbation, ici V.

[38] Bien sûr, l'absence de correction d'énergie au premier ordre est une conséquence du calage additif de $V(\vec{r})$ assurant que $\mathcal{V}(\vec{0})=0$ (voir (30.39)). Comme déjà mentionné, ceci est possible sauf, par exemple, dans le cas exceptionnel du modèle de Kronig - Penney – que l'on sait d'ailleurs résoudre exactement à une dimension (voir problème 17.8.6).

$\mathcal{O}(V)$: tout comme dans le cas 1., ils apparaissent au premier membre de (30.102), multipliés par le crochet de la différence des énergies, qui ne s'annule pas, cependant que le second membre contient des termes d'ordre $\mathcal{O}(V)$ (les g termes du genre $\mathcal{V}c_{\vec{k}-\vec{K}_i}$ – et aussi des termes d'ordre supérieur) et donc s'annule à la limite.

Alors, en extrayant du système (30.95) les seules équations où ces $\vec{K}_i$ figurent au premier membre (il y en a g), il vient :

$$[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)]c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}'\in\tilde{\mathcal{B}},\, \vec{K}'\neq \vec{K}_i} \mathcal{V}(\vec{K}_i - \vec{K}')\, c_{\vec{k}-\vec{K}'} \qquad \forall \vec{K} \in \{\vec{K}_i\} \ . \quad (30.103)$$

Au second membre de (30.103), figurent $\mathrm{g}-1$ coefficients d'ordre zéro, $c_{\vec{k}-\vec{K}_j}$, $j \neq i$, tous les autres étant d'ordre $\mathcal{O}(V)$. Le système (30.103) tronqué aux g lignes pertinentes peut donc s'écrire :

$$[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)]c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}_j\in\{\vec{K}_i\},\, j\neq i} \mathcal{V}(\vec{K}_i - \vec{K}_j)\, c_{\vec{k}-\vec{K}_j} + \mathcal{O}(V^2) \qquad (i = 1,\, 2,\, \ldots,\, \mathrm{g}) \ , \quad (30.104)$$

où tous les coefficients $c_{\vec{k}-\vec{K}_j}$ explicitement écrits sont $\mathcal{O}(V^0)$. Ceci étant, (30.104) montre alors que la correction d'énergie est une somme finie de termes d'ordre $\mathcal{O}(V)$, elle est donc aussi d'ordre $\mathcal{O}(V)$. Si on décide de rester à l'ordre le plus bas non trivial, il suffit de considérer seulement la somme explicitée au second membre de (30.104), et on est réduit à devoir résoudre le système g × g tronqué :

$$\boxed{[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K}_i)]c_{\vec{k}-\vec{K}_i} = \sum_{\vec{K}_j\in\{\vec{K}_i\},\, j\neq i} \mathcal{V}(\vec{K}_i - \vec{K}_j)\, c_{\vec{k}-\vec{K}_j} \quad (i = 1,\, 2,\, \ldots,\, \mathrm{g})} \quad (30.105)$$

La somme au second membre contient au total g − 1 termes ; l'omission des autres termes est le pendant, au *voisinage* d'une dégénérescence, de ce que prescrit la recette de perturbation pour la *stricte* dégénérescence, laquelle précise de ne considérer que la *projection* de l'opérateur de perturbation à l'intérieur de chaque sous-espace dégénéré, en omettant tous les couplages non-diagonaux entre deux sous-espaces dégénérés distincts. En définitive, on obtient un système linéaire g × g pour les g coefficients $c_{\vec{k}-\vec{K}_i}$ relatifs aux g énergies quasi-dégénérées : c'est bien une version plus habile de la méthode de perturbations standard.

Ce *comptage des puissances* est quelque peu laborieux ; une fois les idées bien comprises, on peut reformuler la démarche précédente de façon plus concise en utilisant une technique standard de *partition* à l'aide de projecteurs. Le système de départ, (30.95), peut se récrire en termes de *kets*, compte tenu de l'égalité (30.93) :

$$[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]|f_{\vec{k}-\vec{K}}\rangle = \sum_{\vec{K}'\in\tilde{\mathcal{B}}} \mathcal{V}(\vec{K} - \vec{K}')\, |f_{\vec{k}-\vec{K}'}\rangle \overset{\text{déf}}{=} \hat{\mathcal{V}}|f_{\vec{k}-\vec{K}}\rangle \ , \qquad (30.106)$$

où la dernière égalité sert de définition à l'opérateur noté $\hat{\mathcal{V}}$. Considérons maintenant les g ondes planes $|f_{\vec{k}-\vec{K}_i}\rangle$ dont on a repéré que leurs énergies propres sont toutes voisines

les unes des autres (cas 2.), et définissons les deux projecteurs :

$$P \overset{\text{déf}}{=} \sum_{i=1}^{g} |f_{\vec{k}-\vec{K}_i}\rangle\langle f_{\vec{k}-\vec{K}_i}| \ , \qquad Q \overset{\text{déf}}{=} \mathbf{1} - P \ ; \tag{30.107}$$

on a visiblement $P^2 = P$, d'où $PQ = QP = 0$ et aussi $Q^2 = Q$. En insérant $\mathbf{1} = P + Q$ là où il le faut, le système (30.106) se récrit :

$$(P+Q)[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})](P+Q)|f_{\vec{k}-\vec{K}}\rangle = (P+Q)\hat{\mathcal{V}}(P+Q)|f_{\vec{k}-\vec{K}}\rangle \ ; \tag{30.108}$$

en développant et en remultipliant par P et Q membre à membre, on obtient le système formel[39] :

$$P[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]P|f_{\vec{k}-\vec{K}}\rangle = P\hat{\mathcal{V}}P|f_{\vec{k}-\vec{K}}\rangle + P\hat{\mathcal{V}}Q|f_{\vec{k}-\vec{K}}\rangle \ , \tag{30.110}$$

$$Q[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]Q|f_{\vec{k}-\vec{K}}\rangle = Q\hat{\mathcal{V}}P|f_{\vec{k}-\vec{K}}\rangle + Q\hat{\mathcal{V}}Q|f_{\vec{k}-\vec{K}}\rangle \ . \tag{30.111}$$

De cette équation, on tire :

$$Q|f_{\vec{k}-\vec{K}}\rangle = Q\{Q[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K}) - Q\hat{\mathcal{V}}Q]Q\}^{-1}Q\hat{\mathcal{V}}P|f_{\vec{k}-\vec{K}}\rangle \ , \tag{30.112}$$

que l'on reporte dans (30.110) pour obtenir :

$$P[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]P|f_{\vec{k}-\vec{K}}\rangle = P\hat{\mathcal{V}}P|f_{\vec{k}-\vec{K}}\rangle + \\ P\hat{\mathcal{V}}Q\{Q[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K}) - Q\hat{\mathcal{V}}Q]Q\}^{-1}Q\hat{\mathcal{V}}P|f_{\vec{k}-\vec{K}}\rangle \ , \tag{30.113}$$

En introduisant la résolvante $G_0(\varepsilon) \overset{\text{déf}}{=} (\varepsilon\mathbf{1} - H_0)^{-1}$, et la projection $|f^{(P)}_{\vec{k}-\vec{K}}\rangle$ dans le sous-espace engendré par P, ceci se récrit :

$$[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]|f^{(P)}_{\vec{k}-\vec{K}}\rangle = P\hat{\mathcal{V}}P|f^{(P)}_{\vec{k}-\vec{K}}\rangle + P\hat{\mathcal{V}}QG_0\big(\varepsilon(\vec{k}\,)\big)Q\hat{\mathcal{V}}P|f^{(P)}_{\vec{k}-\vec{K}}\rangle \ ; \tag{30.114}$$

le deuxième terme à droite est visiblement un terme du *second* ordre en V ; on en déduit :

$$[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]|f^{(P)}_{\vec{k}-\vec{K}}\rangle = P\hat{\mathcal{V}}P|f^{(P)}_{\vec{k}-\vec{K}}\rangle + \mathcal{O}(\hat{\mathcal{V}}^2) \ , \tag{30.115}$$

qui est juste la transcription formelle du système $g \times g$ écrit en (30.105), point de départ de l'approximation où l'on ne considère que la *projection* $P\hat{\mathcal{V}}P$ de l'opérateur de perturbation.

Afin de fixer les idées dans ce cas 2., supposons qu'il n'existe que deux valeurs de $\vec{K}$, $\vec{K}_1$ et $\vec{K}_2$, telles que l'égalité $\varepsilon^{(0)}(\vec{k} - \vec{K}_1) = \varepsilon^{(0)}(\vec{k} - \vec{K}_2)$ peut être satisfaite ; comme tout est périodique dans le réseau réciproque, on peut toujours supposer que l'un des deux $\vec{K}_i$ est nul, par exemple $\vec{K}_1$, et poser alors $\vec{K}_2 = \vec{K}$ pour simplifier les

[39]Remarquer que, comme $\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})$ est un scalaire :

$$P[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]Q = Q[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]P = 0 \ . \tag{30.109}$$

écritures. Dès lors, on se trouve dans le cas où $\vec{k}$ est au voisinage d'un seul plan de Bragg, précisément celui de $\vec{K}$. Le système simplifié (30.105), se réduit alors au couple d'équations :

$$[\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k}\,)]\, c_{\vec{k}} - \mathcal{V}(-\vec{K})\, c_{\vec{k}-\vec{K}} = 0 \ , \tag{30.116}$$

$$-\mathcal{V}(\vec{K})\, c_{\vec{k}} + [\varepsilon(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})]\, c_{\vec{k}-\vec{K}} = 0 \ . \tag{30.117}$$

Ce système linéaire homogène n'a de solution non triviale que si son déterminant est nul, ce qui fournit les deux valeurs propres approchées :

$$\varepsilon_{\pm}(\vec{k}\,) = \Sigma\varepsilon(\vec{k}\,) \pm \sqrt{[\Delta\varepsilon(\vec{k}\,)]^2 + |\mathcal{V}(\vec{K})|^2} \equiv \Sigma\varepsilon(\vec{k}\,) \pm h(\vec{k}\,) \ , \tag{30.118}$$

où on a posé :

$$\Sigma\varepsilon(\vec{k}\,) \stackrel{\text{déf}}{=} \frac{1}{2}[\varepsilon^{(0)}(\vec{k}\,) + \varepsilon^{(0)}(\vec{k} - \vec{K})] \ , \qquad \Delta\varepsilon(\vec{k}\,) \stackrel{\text{déf}}{=} \frac{1}{2}[\varepsilon^{(0)}(\vec{k}\,) - \varepsilon^{(0)}(\vec{k} - \vec{K})] \ . \tag{30.119}$$

Cette expression donne la levée de dégénérescence dans le voisinage d'un point de l'espace réciproque où la différence entre $\varepsilon^{(0)}(\vec{k}\,)$ et $\varepsilon^{(0)}(\vec{k}-\vec{K})$ est d'ordre V, c'est-à-dire au voisinage du plan de Bragg de $\vec{K}$ (voir fig. 30.4). Il y a maintenant *deux* solutions distinctes $\varepsilon_{\pm}(\vec{k}\,)$, et qui le restent même quand $\vec{k}$ pointe dans ce plan ; pour cette géométrie, $\Delta\varepsilon = 0$ et $\Sigma\varepsilon(\vec{k}\,) = \varepsilon^{(0)}(\vec{k}\,) = \varepsilon^{(0)}(\vec{k} - \vec{K})$, d'où (voir (30.89)) :

$$\boxed{\varepsilon_{\pm}\big(\vec{k} = \frac{\vec{K}}{2} + \vec{q}_{\perp}\big) = \varepsilon^{(0)}(\vec{k}\,) \pm |\mathcal{V}(\vec{K})|} \tag{30.120}$$

Ainsi, les deux énergies $\varepsilon^{(0)}(\vec{k}\,)$ et $\varepsilon^{(0)}(\vec{k}-\vec{K})$, qui coïncident à l'ordre zéro quand l'extrémité de $\vec{k}$ est dans un plan de Bragg, sont maintenant distinctes : le potentiel périodique a ouvert un *gap* d'énergie, les valeurs propres $\varepsilon_{\pm}(\vec{k}\,)$ différant nettement de $\varepsilon^{(0)}(\vec{k}\,)$. Par continuité, il en va de même tant que $\vec{k}$ reste au voisinage d'un tel plan ; quand l'extrémité de $\vec{k}$ s'éloigne de tout plan de Bragg, les énergies perturbées rejoignent gentiment les $\varepsilon^{(0)}(\vec{k})$, à des termes $\mathcal{O}(V^2)$ près (voir fig. 30.4). Physiquement, le *gap* d'énergie signifie qu'il n'existe aucun état possible pour un électron dans cet intervalle d'énergie.

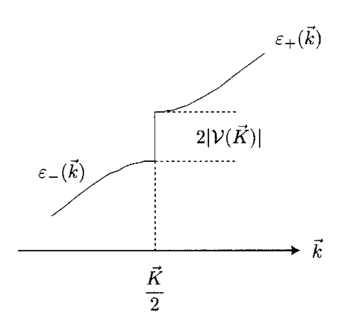


Figure 30.4: Schéma de la levée de dégénérescence pour $\vec{k} = \frac{\vec{K}}{2}$.

En particulier, le plus petit vecteur $\vec{k}$ de cette sorte est $\vec{k} = \frac{\vec{K}}{2}$, parallèle à une direction de $\tilde{\mathcal{B}}$ et on a alors plus précisément :

$$\varepsilon_{\pm}(\vec{k} = \vec{K}/2) = \varepsilon^{(0)}(\vec{k} = \vec{K}/2) \pm |\mathcal{V}(\vec{K})| \ . \tag{30.121}$$

Il existe un vecteur primitif $\vec{b}_0$ de $\tilde{\mathcal{B}}$ tel que $\vec{K} = n\vec{b}_0$ où $n \in \mathbb{Z}$. Si n est impair, le *gap* apparaît en bord de zone ; au contraire, si n est pair, celui-ci se forme en centre de zone. Ainsi, partant de $\vec{k} = 0$, le premier *gap* rencontré dans la direction $\vec{b}_0$ est situé en $\vec{k} = 1 \times \frac{\vec{b}_0}{2}$, il est au bord de BZ1. Le suivant est en $\vec{k} = 2 \times \frac{\vec{b}_0}{2} = \vec{b}_0$; dans le schéma en zones réduites, il est au centre de BZ1, et ainsi de suite.

Le schéma de $\varepsilon^{(0)}(\vec{k})$ en zones réduites montre bien l'apparition d'un *gap* à chaque fois que deux courbes représentant un arc translaté $\varepsilon^{(0)}(\vec{k} - \vec{K})$ se croisent. À l'inverse, si l'on revient à la parabole libre non périodisée (voir fig. 30.5), on voit qu'elle se transforme en une suite d'arcs déformés et disjoints, les (petites) discontinuités survenant pour les valeurs $\vec{k} = \pm\vec{K}/2, \pm\vec{K}, \pm 3\vec{K}/2, \ldots$ L'amplitude des *gaps* décroît quand l'énergie augmente, partant de l'idée de $\mathcal{V}(\vec{q})$ est une fonction décroissante de $\|\vec{q}\|$ (pour $V(\vec{r})$ donné, plus $\vec{q}$ a un grand module, plus l'intégrale de Fourier est oscillante) – d'ailleurs, si l'énergie est très grande, le réseau est *quantitativement* presque invisible pour l'électron. Les *gaps* décroissent lorsque le potentiel V décroît et, à la limite $V = 0$, on retrouve la parabole libre, dénuée de discontinuités.

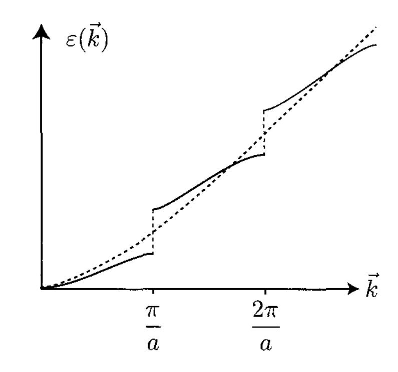


Figure 30.5: Retour à la loi de dispersion non-périodisée : ouverture des *gaps* sur la parabole libre.

Il est utile de regarder d'un peu plus près le voisinage d'un *gap*. Revenons à (30.118) et, pour la commodité, mesurons $\vec{k}$ relativement à $\frac{\vec{K}}{2}$:

$$\vec{k} = \frac{\vec{K}}{2} + \vec{q} \ , \tag{30.122}$$

où $\vec{q}$ est orienté de façon *quelconque* par rapport à $\vec{K}$ (l'extrémité de $\vec{k}$ n'est dans le plan de Bragg que si $\vec{q} \equiv \vec{q}_{\perp} \perp \vec{K}$). Alors $\vec{k} - \vec{K} = -\frac{1}{2}\vec{K} + \vec{q}$ et :

$$\Sigma\varepsilon = \varepsilon^{(0)}\big(\frac{\vec{K}}{2}\big) + \varepsilon^{(0)}(\vec{q}) \ , \qquad \Delta\varepsilon = \frac{\hbar^2}{m}\,\vec{q}.\vec{K} \ . \tag{30.123}$$

Les valeurs propres approchées données par (30.118) s'écrivent maintenant :

$$\varepsilon_{\pm}(\vec{k}\,) = \varepsilon^{(0)}(\frac{\vec{K}}{2}) + \varepsilon^{(0)}(\vec{q}\,) \pm \sqrt{(\frac{\hbar^2}{2m})^2(\vec{K}.\vec{q}\,)^2 + |\mathcal{V}(\vec{K})|^2} \ . \tag{30.124}$$

Quand de surcroît l'extrémité de $\vec{k}$ est dans le plan de Bragg, il reste :

$$\varepsilon_{\pm}(\vec{k} = \frac{\vec{K}}{2} + \vec{q}_{\perp}) = \varepsilon^{(0)}(\frac{\vec{K}}{2}) + \varepsilon^{(0)}(\vec{q}_{\perp}\,) \pm |\mathcal{V}(\vec{K})| = \frac{\hbar^2}{2m}(\vec{k}^2 - \vec{k}.\vec{K} + \frac{\vec{K}^2}{2}) \pm |\mathcal{V}(\vec{K})| \ , \tag{30.125}$$

en agrément avec (30.120). De ceci il résulte :

$$\left(\vec{\nabla}_{\vec{k}}\varepsilon_{\pm}\right)_{\vec{k}=\frac{\vec{K}}{2}+\vec{q}_{\perp}} = \frac{\hbar^2}{m}(\vec{k} - \frac{\vec{K}}{2}) \equiv \frac{\hbar^2}{m}\vec{q}_{\perp} \ , \tag{30.126}$$

quantité qui s'annule bien en $\vec{k} = \frac{\vec{K}}{2}$ (voir (30.75)). Le gradient de $\varepsilon_{\pm}(\vec{k}\,)$ donnant la direction de la normale à la surface $\varepsilon_{\pm}(\vec{k}\,) = \mathrm{C^{ste}}$, on voit qu'une surface d'énergie constante intersecte perpendiculairement tout plan de Bragg.

Par ailleurs, quand $\vec{k}$ est perpendiculaire à $\mathcal{P}_{\vec{K}/2}$ (c'est-à-dire $\vec{k} \parallel \vec{K}$ et $\vec{q} \parallel \vec{K}$) :

$$\varepsilon_{\pm}(\vec{k} \parallel \vec{K}) = \varepsilon^{(0)}(\frac{\vec{K}}{2}) + \varepsilon^{(0)}(\vec{q}\,) \pm \sqrt{(\frac{\hbar^2 Kq}{2m})^2 + |\mathcal{V}(\vec{K})|^2} \ . \tag{30.127}$$

Cette expression permet de visualiser la variation des énergies de bande quand $\vec{k}$ traverse un plan de Bragg tout en restant paralèlle à $\vec{K}$; quand $\vec{k}$ est près de $\frac{\vec{K}}{2}$, $\vec{q}$ est très petit et on peut faire un développement limité en q, et on trouve :

$$\varepsilon_{\pm}(\vec{k} \parallel \vec{K},\, \vec{k} \simeq \frac{\vec{K}}{2}) \simeq \varepsilon^{(0)}(\frac{\vec{K}}{2}) \pm |\mathcal{V}(\vec{K})| + \frac{\hbar^2 q^2}{2m}\Big(1 \pm \frac{\hbar^2 K^2}{4m|\mathcal{V}(\vec{K})|}\Big) \ , \tag{30.128}$$

ce développement n'ayant de sens que si $q \ll \frac{2m|\mathcal{V}(\vec{K})|}{\hbar^2 K}$ (ce qui interdit de prendre la limite de potentiel nul dans les expressions qui suivent). Depuis le début, on fait l'hypothèse que le couplage de réseau est faible devant les autres énergies pertinentes ; se donnant l'ordre de grandeur a du pas du réseau, la seule autre énergie disponible (pour un électron) est $\frac{\hbar^2}{ma^2}$; compte tenu de (30.91), le 1 dans la parenthèse en facteur de q^2 dans (30.128) est très petit devant la fraction ; le signe du coefficient de q^2 est positif pour ε_+, négatif pour ε_- :

$$\boxed{\varepsilon_{\pm}(\vec{k} \parallel \vec{K},\, \vec{k} \simeq \frac{\vec{K}}{2}) \simeq \varepsilon^{(0)}(\frac{\vec{K}}{2}) \pm |\mathcal{V}(\vec{K})| \pm \frac{\hbar^2}{2m}\Big(\frac{\hbar^2 K^2}{4m|\mathcal{V}(\vec{K})|} \mp 1\Big)\, q^2} \tag{30.129}$$

Ceci montre que, comme attendu, les deux bandes se repoussent et ressemblent à deux paraboles de concavités opposées, séparées, là où elles sont extrémales, par l'énergie $2|\mathcal{V}(\vec{K})|$, qui est l'énergie de *gap*.

Au voisinage du *gap*, et par référence avec la loi de dispersion libre (30.46), tout se passe comme si l'électron situé dans la bande ε_- avait une masse *négative*. De fait, on définit une masse effective m_{eff} comme suit :

$$m_{\text{eff}\,\pm} = \hbar^{-2}\left(\frac{\partial^2\varepsilon}{\partial q^2}\right)^{-1} = \pm\frac{m}{\frac{\hbar^2K^2}{4m|\mathcal{V}(\vec{K})|}\mp 1} \simeq \pm\frac{4m|\mathcal{V}(\vec{K})|}{\hbar^2K^2}\,m \ , \tag{30.130}$$

où m est la masse *nue* de l'électron. Cette masse est positive pour la bande supérieure et *négative* pour la bande inférieure : à chaque fois, le signe de la masse est celui de la concavité des bandes. Géométriquement, on voit que la masse effective est (en module) très petite devant la masse nue, ce que confirme l'expression (30.130) :

$$\boxed{|m_{\text{eff}\,\pm}| \ll m} \tag{30.131}$$

Par exemple avec $|\mathcal{V}(K)| \sim 0,1$ eV, on a typiquement $|m_{\text{eff}\,\pm}| \sim 10^{-2}\,m$: nul doute que cette petite masse effective jouera un rôle important pour les propriétés de transport.

Le résultat élémentaire (30.130) vaut lorsque $\vec{k}$ est parallèle à $\vec{K}$; plus généralement, on définit un tenseur de masse effective d'éléments :

$$\boxed{m_{\text{eff}\,uv} \stackrel{\text{déf}}{=} \hbar^{-2}\left(\frac{\partial^2\varepsilon}{\partial k_u\partial k_v}\right)^{-1} \qquad (u,\, v = x,\, y,\, z)} \tag{30.132}$$

qui joue un rôle essentiel dans la description du transport électronique dans un métal ; en particulier, de fortes anisotropies de la masse effective produisent une conductivité très anisotrope.

Une fois les valeurs propres (approchées) obtenues, il est possible de trouver les fonctions propres. Pour simplifier les notations, on suppose dans la suite que le cristal possède la symétrie d'inversion[40] $V(-\vec{r}) = V(\vec{r})$, ce qui, allié au fait que V est réel, assure que $\mathcal{V}(\vec{K}) \in \mathbb{R}$. Les deux équations (30.116) et (30.117) peuvent être considérées comme la forme matricielle de l'équation aux vecteurs propres :

$$H_{\text{eff}}\,|\psi_{\vec{k}}\rangle = \varepsilon(\vec{k}\,)|\psi_{\vec{k}}\rangle \ , \qquad H_{\text{eff}} \stackrel{\text{déf}}{=} \Sigma\varepsilon\,\mathbf{1} + \Delta\varepsilon\,\sigma_z + \mathcal{V}(\vec{K})\,\sigma_x \ , \tag{30.133}$$

définie dans le sous-espace $\{|f_{\vec{k}}\rangle,\, |f_{\vec{k}-\vec{K}}\rangle\}$, avec $\langle\vec{r}\,|f_{\vec{q}}\rangle \equiv f_{\vec{q}}(\vec{r}\,) = \mathrm{e}^{\mathrm{i}\vec{q}.\vec{r}}$. On a :

$$H_{\text{eff}} - \Sigma\varepsilon\,\mathbf{1} = h(\vec{k}\,)\,(\cos 2\theta\sigma_z + \sin 2\theta\sigma_x) = h(\vec{k}\,)\,\mathrm{e}^{-\mathrm{i}\theta\sigma_y}\,\sigma_z\mathrm{e}^{+\mathrm{i}\theta\sigma_y} \ . \tag{30.134}$$

où $h(\vec{k}\,) \stackrel{\text{déf}}{=} \sqrt{[\Delta\varepsilon]^2 + |\mathcal{V}(\vec{K})|^2}$ (voir (30.118)) et avec :

$$\boxed{\tan 2\theta = \frac{\mathcal{V}(\vec{K})}{\Delta\varepsilon}} \tag{30.135}$$

[40] C'est toujours vrai pour une structure cristalline sans motif, avec des objets à symétrie sphérique.

Les deux vecteurs propres $|\psi_{\vec{k}\,\pm}\rangle$ sont alors égaux à :

$$|\psi_{\vec{k}\,+}\rangle = \mathrm{e}^{-\mathrm{i}\theta\sigma_y}\,|f_{\vec{k}}\rangle \ , \qquad |\psi_{\vec{k}\,-}\rangle = \mathrm{e}^{-\mathrm{i}\theta\sigma_y}\,|f_{\vec{k}-\vec{K}}\rangle \ . \tag{30.136}$$

En effet, H_{eff} agissant sur $|\psi_{\vec{k}\,+}\rangle$ donne :

$$\begin{aligned} H_{\mathrm{eff}}|\psi_{\vec{k}\,+}\rangle &= h(\vec{k}\,)\mathrm{e}^{-\mathrm{i}\theta\sigma_y}\sigma_z\mathrm{e}^{+\mathrm{i}\theta\sigma_y}\mathrm{e}^{-\mathrm{i}\theta\sigma_y}|f_{\vec{k}}\rangle = h(\vec{k}\,)\mathrm{e}^{-\mathrm{i}\theta\sigma_y}\sigma_z|f_{\vec{k}}\rangle \\ &= +h(\vec{k}\,)\mathrm{e}^{-\mathrm{i}\theta\sigma_y}|f_{\vec{k}}\rangle \equiv h(\vec{k}\,)\,|\psi_{\vec{k}\,+}\rangle \ , \end{aligned} \tag{30.137}$$

et de même :

$$\begin{aligned} H_{\mathrm{eff}}\,|\psi_{\vec{k}\,-}\rangle &= h(\vec{k}\,)\,\mathrm{e}^{-\mathrm{i}\theta\sigma_y}\sigma_z\,\mathrm{e}^{+\mathrm{i}\theta\sigma_y}\mathrm{e}^{-\mathrm{i}\theta\sigma_y}\,|f_{\vec{k}-\vec{K}}\rangle = h(\vec{k}\,)\,\mathrm{e}^{-\mathrm{i}\theta\sigma_y}\sigma_z\,|f_{\vec{k}-\vec{K}}\rangle\rangle \\ &= -\,h(\vec{k}\,)\,\mathrm{e}^{-\mathrm{i}\theta\sigma_y}\,|f_{\vec{k}-\vec{K}}\rangle \equiv\, -\,h(\vec{k}\,)\,|\psi_{\vec{k}\,-}\rangle \ . \end{aligned} \tag{30.138}$$

De (30.136) on déduit la forme explicite des fonctions de Bloch approchées $\psi_{\vec{k}\,\pm}$:

$$\psi_{\vec{k}+}(\vec{r}\,) = \cos\theta f_{\vec{k}}(\vec{r}\,){+}\sin\theta f_{\vec{k}-\vec{K}}(\vec{r}\,) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}(\cos\theta{+}\sin\theta\,\mathrm{e}^{-\mathrm{i}\vec{K}.\vec{r}}) \equiv \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}u_{\vec{k}+}(\vec{r}\,) \ , \tag{30.139}$$

$$\psi_{\vec{k}-}(\vec{r}\,) = -\sin\theta f_{\vec{k}}(\vec{r}) + \cos\theta f_{\vec{k}-\vec{K}}(\vec{r}\,) = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}(-\sin\theta + \cos\theta\,\mathrm{e}^{-\mathrm{i}\vec{K}.\vec{r}}) \equiv \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}}u_{\vec{k}-}(\vec{r}\,) \ . \tag{30.140}$$

Il est visible que les $\psi_{\vec{k}\,\pm}(\vec{r}\,)$, (30.139) et (30.140), satisfont le théorème de Bloch, les $u_{\vec{k}\,\pm}$ étant manifestement périodiques dans $\mathcal{B}$.

Chacune de ces fonctions redonne une simple onde plane dès que $\theta \simeq 0$; ceci arrive dès que $|\vec{K}.(\vec{k} - \vec{K}/2)|$ est supérieur ou de l'ordre de $m\hbar^{-2}|\mathcal{V}(\vec{K})| \ll a^{-2}$: dès que l'extrémité de $\vec{k}$ s'éloigne un tout petit peu du plan de Bragg, les fonctions perturbées coïncident pratiquement avec les ondes planes, propres à l'ordre zéro. D'une façon générale, on trouve que $|\psi_{\vec{k}\,\pm}(\vec{r}\,)|^2 = 1 \pm \sin 2\theta\, \cos\vec{K}.\vec{r}$, qui reprend la même valeur si on translate $\vec{r}$ de n'importe quel $\vec{R} \in \mathcal{B}$.

Les expressions (30.139) et (30.140) se simplifient lorsque l'extrémité de $\vec{k}$ est dans un plan de Bragg. Alors, $\Delta\varepsilon = 0$, $\theta = +\frac{\pi}{4}$ si $\mathcal{V}(\vec{K})$ est positif (voir (30.135)) et :

$$\psi_{\vec{k}\,+}(\vec{r}\,) = \sqrt{2}\,\mathrm{e}^{\mathrm{i}(\vec{k}-\frac{\vec{K}}{2}).\vec{r}}\cos\frac{\vec{K}.\vec{r}}{2} \ , \quad \psi_{\vec{k}\,-}(\vec{r}\,) = -\mathrm{i}\sqrt{2}\,\mathrm{e}^{\mathrm{i}(\vec{k}-\frac{\vec{K}}{2}).\vec{r}}\sin\frac{\vec{K}.\vec{r}}{2} \tag{30.141}$$

Maintenant, les fonctions propres ne sont plus des ondes planes, mais (à un facteur près) des ondes *stationnaires*, combinaisons linéaires à poids égaux des deux exponentielles $\mathrm{e}^{\pm\mathrm{i}\vec{K}.\vec{r}/2}$: c'est pourquoi il est usuel de dire, en pareil cas, que l'électron subit une *réflexion de Bragg* (voir problème (30.7.13) pour un phénomène équivalent avec des *atomes*).

30.4.2 L'approximation des liaisons fortes

Il s'agit de la situation extrême opposée à celle que l'on vient de décrire : on va maintenant supposer que le potentiel ressenti par un électron au voisinage d'un cœur est

essentiellement celui de ce dernier, le cristal pouvant alors être vu, en première approximation, comme une simple juxtaposition d'atomes, objets neutres en interaction faible. Il en résulte que les *bonnes* fonctions d'ordre zéro sont maintenant les fonctions atomiques centrées sur un atome donné, tout le contraire d'ondes planes infiniment étendues. Ces orbitales atomiques seront notées $\chi_{\nu\vec{R}}(\vec{r})$, ν désignant un ensemble de nombres quantiques définis au niveau de l'atome isolé. En gros, l'approximation consiste à supposer que, quand on s'éloigne d'un cœur, le potentiel de réseau $V(\vec{r})$, commence à différer du pur potentiel atomique quand on atteint une distance au moins égale à la portée des fonctions atomiques $\chi_{\nu\vec{R}}(\vec{r})$.

Tous les sites du réseau étant équivalents, toutes les fonctions $\chi_{\nu\vec{R}}$, ν fixé, sont en fait les répliques les unes des autres et se transforment les unes dans les autres par les translations de $\mathcal{B}$. Si on note simplement par $\chi_\nu(\vec{r}) \equiv \chi_{\nu\,\vec{R}=\vec{0}}(\vec{r})$ les fonctions centrées sur le nœud origine, on a :

$$\boxed{\chi_{\nu\vec{R}}(\vec{r}) = \mathcal{T}(\vec{R})\,\chi_\nu(\vec{r}) = \chi_\nu(\vec{r}-\vec{R})} \tag{30.142}$$

À ce stade, on dispose *a priori*, d'un jeu constitué par tous les états propres d'un Hamiltonien atomique (celui de l'atome situé au nœud origine $\vec{R} = \vec{0}$, H_0) et de toutes leurs répliques par le jeu des translations du réseau de Bravais. Cet ensemble est visiblement surcomplet, au sens où ses éléments ne sont pas tous linéairement indépendants les uns des autres. En effet, H_0 à lui seul offre une base complète ; en conséquence, toute fonction peut être développée sur cette base, en particulier n'importe quelle fonction du type $\chi_{\nu\,\vec{R}\neq\vec{0}}(\vec{r})$. Il existe donc des développements du genre :

$$\chi_{\nu\vec{R}\neq\vec{0}}(\vec{r}) = \sum_{\nu'} C_{\nu\nu'}(\vec{r})\,\chi_{\nu'}(\vec{r}) \ , \tag{30.143}$$

avec leurs coefficients non tous identiquement nuls. Il en résulte que l'on ne doit pas retenir pour $\{\chi_\nu\}_\nu$ *tous* les états propres de H_0, mais seulement un sous-ensemble. En pratique, d'ailleurs, on ne peut jamais introduire (pour des raisons techniques) qu'un petit nombre de fonctions de ce genre. En outre, l'idée physique sous-jacente à l'approximation en cours repose sur la notion d'état localisé : de ce point de vue, il serait incongru d'inclure parmi les χ_ν les états non liés de H_0 – qui sont pourtant nécessaires pour assurer la complétude de la base propre de cet opérateur[41]. Évacuer d'emblée ces états mal venus élimine de fait la difficulté liée à la surcomplétude.

D'ailleurs, finalement, rien n'oblige à choisir comme états localisés certains états propres de H_0. Le caractère essentiel à préserver est la localisation autour de $\vec{R} = 0$, ce qui reste acquis quand on fait des combinaisons linéaires d'états $\chi_\nu(\vec{r})$: une combinaison linéaire en nombre fini d'états localisés donne clairement une fonction qui l'est encore. En définitive, l'ingrédient de base dans la suite est une (en fait plusieurs) fonction(s)

[41] L'inclusion de tels états délocalisés (étendus) ferait repartir clairement en direction de l'approximation des électrons presque libres.

notée(s) $\phi(\vec{r})$, localisée(s) dans le voisinage de $\vec{R} = \vec{0}$ et (mentalement) développable(s) sur un *petit* jeu de g fonctions atomiques χ_ν localisées sur ce site :

$$\phi(\vec{r}) = \sum_{\nu=1}^{\mathrm{g}} c_\nu \, \chi_\nu(\vec{r}) \ , \tag{30.144}$$

avec :

$$H_0 \chi_\nu(\vec{r}) = E_\nu \chi_\nu(\vec{r}) \ . \tag{30.145}$$

Les coefficients c_ν dans (30.144) sont déterminés suivant une procédure qu'il n'est pas utile de préciser ici et que l'on considèrera comme donnés. Celle-ci étant achevée, on dispose de g fonctions ϕ localisées autour du site $\vec{R} = 0$. Un label permet de les distinguer les unes des autres, le cas échéant, mais est omis dans la suite sauf nécessité. Propres du Hamiltonien atomique relatif à ce site, les énergies propres correspondantes sont bien séparées, leurs différences étant de l'ordre de l'eV.

Dans l'hypothèse des liaisons fortes, l'image physique à l'ordre zéro est celle d'un électron confiné près d'un site et placé dans une orbitale $\chi_{\nu\vec{R}}$ (ou une fonction du type ϕ ci-dessus), fortement localisée. $\vec{R}$ variant, toutes ces fonctions sont, à cet ordre, strictement équivalentes : on peut placer l'électron n'importe où, son énergie (à cet ordre) sera toujours la même[42]. En réalité, l'électron a la possibilité de sauter d'un site à l'autre, ce qui se traduit par le fait que les éléments de matrice $\langle \chi_{\nu\vec{R}} | H | \chi_{\nu'\vec{R}'} \rangle$ sont différents de zéro – quoique petits pour que le schéma de liaisons fortes ait un sens. L'existence de ce mouvement par sauts donne à l'électron une énergie cinétique, qui se traduit par une loi de dispersion $\varepsilon(\vec{k})$, exprimant comment varie son énergie en fonction de sa pseudo-impulsion. Comme l'amplitude des sauts est réputée petite, la variation de ε en $\vec{k}$ sera faible : au total, la bande sera *peu dispersée* (à l'ordre zéro, elle est strictement *plate* : l'énergie d'un atome isolé est ce qu'elle est, et ne saurait dépendre de $\vec{k}$).

Outre $\phi(\vec{r})$, il convient donc d'introduire toutes les fonctions translatées :

$$\boxed{\phi_{\vec{R}}(\vec{r}) \stackrel{\text{déf}}{=} \mathcal{T}(\vec{R})\phi(\vec{r}) = \phi(\vec{r} - \vec{R})} \tag{30.146}$$

notons dès maintenant que rien n'impose aux $\phi_{\vec{R}}$ d'être orthogonales entre elles ; en général :

$$\langle \phi_{\vec{R}} | \phi_{\vec{R}'} \rangle \neq 0 \quad \forall \vec{R}, \, \vec{R}' \ . \tag{30.147}$$

Quelle que soit l'image en arrière-plan (suggérant une approximation ou une autre), il s'agit toujours de trouver les états d'un électron dans un potentiel périodique $V(\vec{r})$, qui, de ce fait, satisfont le théorème de Bloch. À partir des $\phi_{\vec{R}}(\vec{r})$, on doit trouver des états $\psi_{\vec{k}}(\vec{r})$ également propres de tous les $\mathcal{T}(\vec{R})$, $\mathcal{T}(\vec{R})\psi_{\vec{k}}(\vec{r}) = \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}}\psi_{\vec{k}}(\vec{r})$, égalité qui, en vertu de la définition (30.8), est équivalente à :

$$\psi_{\vec{k}}(\vec{r} + \vec{R}) = \mathrm{e}^{+\mathrm{i}\vec{k}.\vec{R}}\psi_{\vec{k}}(\vec{r}) \ . \tag{30.148}$$

[42]Le problème est donc très fortement dégénéré.

Il existe des combinaisons linéaires des $\phi_{\vec{R}}(\vec{r})$ satisfaisant la relation (30.148). En effet, posons :

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}')\phi_{\vec{R}'}(\vec{r}) = \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}')\phi(\vec{r} - \vec{R}') \ ; \tag{30.149}$$

reportant dans (30.148), il vient :

$$\sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}')\phi(\vec{r} + \vec{R} - \vec{R}') = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}')\phi(\vec{r} - \vec{R}') \ ; \tag{30.150}$$

posant $\vec{R}'' = \vec{R}' - \vec{R}$ au premier membre :

$$\sum_{\vec{R}'' \in \mathcal{B}} A_{\vec{k}}(\vec{R} + \vec{R}'')\phi(\vec{r} - \vec{R}'') = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \sum_{\vec{R}' \in \mathcal{B}} A_{\vec{k}}(\vec{R}')\phi(\vec{r} - \vec{R}') \ . \tag{30.151}$$

Les variables $\vec{R}'$ et $\vec{R}''$ étant muettes, on en déduit $A_{\vec{k}}(\vec{R} + \vec{R}') = \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} A_{\vec{k}}(\vec{R}')$, ce qui fournit la solution :

$$A_{\vec{k}}(\vec{R}) = C\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \qquad (C = \text{ constante arbitraire}) \ . \tag{30.152}$$

Ainsi, étant donné un jeu de fonctions $\phi_{\vec{R}}$ déduites les unes des autres par les translations de $\mathcal{B}$, toute fonction du type[43] :

$$\psi_{\vec{k}}(\vec{r}) = C \sum_{\vec{R} \in \mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \phi_{\vec{R}}(\vec{r}) \tag{30.154}$$

satisfait le théorème de Bloch, la constante C devant être choisie par normalisation. Avec des conditions cycliques dans une grande boîte contenant N sites au total, on a :

$$\langle \psi_{\vec{k}} | \psi_{\vec{k}} \rangle = |C|^2 \sum_{\vec{R} \in \mathcal{B}} \sum_{\vec{R}' \in \mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.(\vec{R}' - \vec{R})} \langle \phi_{\vec{R}} | \phi_{\vec{R}'} \rangle = |C|^2 \sum_{\vec{R} \in \mathcal{B}} \sum_{\vec{R}'' \in \mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}''} \langle \phi_{\vec{0}} | \phi_{\vec{R}''} \rangle \tag{30.155}$$

[43] Noter toutefois que ce résultat ne dépend nullement de l'aspect localisé éventuel des $\phi_{\vec{R}}(\vec{r})$: le point essentiel est que toutes ces fonctions se déduisent les unes des autres par translation.

Il est aussi intéressant de remarquer que la fonction de bonne symétrie (30.154) peut être obtenue directement en faisant agir l'opérateur $\mathbf{T}$:

$$\mathbf{T} = \sum_{\vec{R} \in \mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \mathcal{T}(\vec{R}) \tag{30.153}$$

sur une fonction $\phi_{\vec{R}_0}$ quelconque vis-à-vis de la symétrie de translation. Ce fait est général : quelle que soit la nature de la symétrie (géométrique, de permutations, etc.), la théorie des groupes permet de définir les bons projecteurs construisant automatiquement une fonction de symétrie requise. Ces opérateurs sont des combinaisons linéaires des opérateurs du groupe de symétrie, les coefficients étant (ici) les valeurs propres de ces opérateurs (plus généralement, ce sont les *caractères* des opérateurs, définis relativement à un type de symétrie donné). Pour un ensemble de particules identiques, on a vu que la construction de la fonction ayant la bonne symétrie de permutation relève exactement du même traitement : les opérateurs Y_{I} définis en (22.44) et égaux à $N!^{-1} \sum_\lambda \varepsilon_\lambda P_\lambda$ agissant sur une fonction de symétrie *quelconque* produisent automatiquement la fonction de symétrie voulue.

Enfin, on note la similitude entre l'expression (30.154) des fonctions $\psi_{\vec{k}}(\vec{r})$ et celle des orbitales moléculaires du benzène, (chapitre 28, éq. (28.91)). Ici, c'est le groupe des translations qui est à l'œuvre, là-bas c'est le groupe des rotations, chaque groupe injectant un facteur de phase spécifique dans le développement des états propres.

où les produits scalaires $\langle \phi_{\vec{R}} | \phi_{\vec{R}'} \rangle$, $\vec{R} \neq \vec{R}'$, *ne sont pas nuls* en général. Comme la somme interne sur $\vec{R}''$ est indépendante de $\vec{R}$, la somme sur $\vec{R}$ donne un simple facteur N, d'où la relation de normalisation fixant la constante multiplicative C :

$$\boxed{N\,|C|^2 \sum_{\vec{R}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\,\langle \phi_{\vec{R}=\vec{0}} | \phi_{\vec{R}} \rangle = 1} \tag{30.156}$$

Si l'on dispose de g fonctions localisées de type ϕ, $\{\phi_n\}$, et donc aussi de toutes leurs translatées $\{\phi_{n\vec{R}}\}$, on pourra construire autant de fonctions de Bloch à $\vec{k}$ donné :

$$\psi_{n\vec{k}}(\vec{r}) = C \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\,\phi_{n\vec{R}}(\vec{r}) \tag{30.157}$$

Par référence avec le premier énoncé du théorème de Bloch, la fonction $u_{n\vec{k}}(\vec{r})$ est :

$$u_{n\vec{k}}(\vec{r}) = C \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.(\vec{R}-\vec{r})}\,\phi_{n\vec{R}}(\vec{r}) \equiv C \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.(\vec{R}-\vec{r})}\,\phi_{n\vec{0}}(\vec{r}-\vec{R}) \ , \tag{30.158}$$

fonction qui est visiblement périodique dans $\mathcal{B}$. Noter que si les fonctions $\psi_{n\vec{k}}$ satisfont le théorème de Bloch par construction, elles ne sont pas pour autant des états de Bloch *exacts*, au sens où la fonction $u_{n\vec{k}}$ doit satisfaire sa propre équation, ce qui n'est pas assuré en général avec le choix limité des fonctions atomiques $\phi_{n\vec{R}}$: ce n'est pas parce qu'une fonction a la bonne symétrie qu'elle est fonction propre du Hamiltonien (comparer avec (30.181)) – voir aussi le commentaire introduisant la relation logique (30.24).

L'usage des fonctions $\psi_{n\vec{k}}$ simplifie considérablement le problème sur un plan technique, puisqu'elles prédiagonalisent H en conséquence du fait que, satisfaisant le théorème de Bloch, elles sont par construction fonctions propres des opérateurs de translation :

$$\mathcal{T}(\vec{R})\,\psi_{n\vec{k}}(\vec{r}) = \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}}\,\psi_{n\vec{k}}(\vec{r}) \ . \tag{30.159}$$

En effet, H commutant avec tous les opérateurs $\mathcal{T}(\vec{R})$, $H\mathcal{T}(\vec{R}) = \mathcal{T}(\vec{R})H$; prenons l'élément de matrice de cette égalité entre deux fonctions $\psi_{n\vec{k}}$ et $\psi_{n'\vec{k}'}$:

$$\langle \psi_{n\vec{k}} | H\mathcal{T}(\vec{R}) | \psi_{n'\vec{k}'} \rangle = \langle \psi_{n\vec{k}} | \mathcal{T}(\vec{R}) H | \psi_{n'\vec{k}'} \rangle \ ; \tag{30.160}$$

en vertu de (30.159), le premier membre est :

$$\mathrm{e}^{-\mathrm{i}\vec{k}'.\vec{R}}\,\langle \psi_{n\vec{k}} | H | \psi_{n'\vec{k}'} \rangle \ . \tag{30.161}$$

Par définition de l'adjoint d'un opérateur, le second membre de (30.160) est :

$$\big(|\psi_{n\vec{k}}\rangle,\, \mathcal{T}(\vec{R})H|\psi_{n'\vec{k}'}\rangle\big) = \big([\mathcal{T}(\vec{R})]^\dagger |\psi_{n\vec{k}}\rangle,\, H|\psi_{n'\vec{k}'}\rangle\big) \ , \tag{30.162}$$

et comme les $\mathcal{T}$ sont unitaires et que $[\mathcal{T}(\vec{R})]^{-1} = \mathcal{T}(-\vec{R})$, ceci vaut encore :

$$\big(\mathcal{T}(-\vec{R})|\psi_{n\vec{k}}\rangle,\, H|\psi_{n'\vec{k}'}\rangle\big) = \big(\mathrm{e}^{-\mathrm{i}\vec{k}.(-\vec{R})}|\psi_{n\vec{k}}\rangle,\, H|\psi_{n'\vec{k}'}\rangle\big) = \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}}\,\big(|\psi_{n\vec{k}}\rangle,\, H|\psi_{n'\vec{k}'}\rangle\big) \ , \tag{30.163}$$

où l'antilinéarité dans le *bra* a été utilisée pour écrire la dernière égalité. La quantité à droite n'est autre que $\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}}\,\langle\psi_{n\vec{k}}|H|\psi_{n'\vec{k}'}\rangle$. En rapprochant maintenant (30.161) et (30.163), il vient :

$$(\mathrm{e}^{-\mathrm{i}\vec{k}'.\vec{R}} - \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}})\,\langle\psi_{n\vec{k}}|H|\psi_{n'\vec{k}'}\rangle = 0 \ , \tag{30.164}$$

d'où il résulte que si $\vec{k}' \neq \vec{k}$, l'élément de matrice de H est nul : les $\psi_{n\vec{k}}$ assurent bien une prédiagonalisation de H. Toutefois, H possède *a priori* des éléments de matrice non-nuls entre deux états de même $\vec{k}$ mais de n différents :

$$\langle\psi_{n\vec{k}}|H|\psi_{n'\vec{k}}\rangle \neq 0 \ ; \tag{30.165}$$

cependant, et c'est l'esprit même de l'approximation, ces éléments de matrice n'apportent que des petites corrections quand les énergies atomiques sont bien séparées les unes des autres. Quoi qu'il en soit, à ce stade, le problème est posé dans de bonnes conditions et on est réduit à résoudre une équation aux valeurs propres dans un espace de dimension g. Cet espace étant paramétré par $\vec{k}$, les g valeurs propres seront des fonctions de $\vec{k}$: ce sont les g relations de dispersion cherchées, $\varepsilon_m(\vec{k}\,)$.

Afin d'illustrer les idées essentielles, supposons qu'en fait *une seule* orbitale atomique est introduite sur chaque site (g = 1), $\chi_{1\vec{R}}$, réplique par translation de l'orbitale χ_1 définie sur le site à l'origine. Les définitions précédentes restent valables, avec partout g = 1 ; l'indice ν ne prend désormais qu'une seule valeur $\nu = 1$ mais est maintenu pour des raisons de clarté dans les notations. Notons aussi que la fonction ϕ définie en (30.144) se réduit à χ_1 :

$$\phi(\vec{r}\,) = \chi_1(\vec{r}\,) \ , \qquad H_0\chi_1 = E_1\chi_1 \ . \tag{30.166}$$

Ce cadre *minimal* étant admis, et en conséquence de (30.164), H se trouve représenté par une matrice complètement diagonale sur les N fonctions $\{\psi_{\vec{k}}\}_{\vec{k}}$. Il reste dès lors à calculer l'élément diagonal $\langle\psi_{\vec{k}}|H|\psi_{\vec{k}}\rangle$, qui n'est autre que la valeur propre approchée $\varepsilon(\vec{k}\,)$. Pour calculer cet élément de matrice, on part de :

$$\psi_{\vec{k}} = C\sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\,\chi_1(\vec{r}-\vec{R}\,) \ , \tag{30.167}$$

d'où :

$$\varepsilon(\vec{k}\,) = |C|^2\sum_{\vec{R}}\sum_{\vec{R}'} \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}'}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\langle\chi_{1\vec{R}'}|H|\chi_{1\vec{R}}\rangle \ ; \tag{30.168}$$

l'élément de matrice se transforme en jouant avec l'invariance par translation de H :

$$\langle\chi_{1\vec{R}'}|H|\chi_{1\vec{R}}\rangle = \int \chi_1^*(\vec{r}-\vec{R}')H\chi_1(\vec{r}-\vec{R}\,)\,\mathrm{d}^3r = \int \chi_1^*(\vec{r}\,')H\chi_1(\vec{r}\,'-\vec{R}+\vec{R}')\,\mathrm{d}^3r' \ ; \tag{30.169}$$

l'intégrale à droite est $\langle\chi_1|H|\chi_{1\vec{R}-\vec{R}'}\rangle$, d'où :

$$\varepsilon(\vec{k}\,) = |C|^2\sum_{\vec{R}}\sum_{\vec{R}'} \mathrm{e}^{\mathrm{i}\vec{k}.(\vec{R}-\vec{R}')}\langle\chi_1|H|\chi_{1\vec{R}-\vec{R}'}\rangle = |C|^2\sum_{\vec{R}}\sum_{\vec{R}''} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}''}\langle\chi_1|H|\chi_{1\vec{R}''}\rangle \ . \tag{30.170}$$

À nouveau, et compte tenu de l'invariance de H par translation, la sommation sur $\vec{R}$ fournit N fois le même terme, de sorte que :

$$\varepsilon(\vec{k}) = N|C|^2 \sum_{\vec{R}'} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}'} \langle \chi_1 | H | \chi_{1\vec{R}'} \rangle \ . \tag{30.171}$$

Pour achever le calcul, on effectue la partition suivante du Hamiltonien. Le Hamiltonien d'un électron dans le réseau est :

$$H = \frac{\vec{p}^2}{2m} + \sum_{\vec{R}\in\mathcal{B}} v(\vec{r} - \vec{R}) \ , \tag{30.172}$$

où $v(\vec{r})$ est l'énergie potentielle d'un électron soumis à l'attraction du seul nœud situé à l'origine ; en isolant le terme $\vec{R} = 0$:

$$H = \left[\frac{\vec{p}^2}{2m} + v(\vec{r})\right] + \sum_{\vec{R}\in\mathcal{B}^*} v(\vec{r} - \vec{R}) \equiv H_0 + \Delta V \ , \tag{30.173}$$

$\mathcal{B}^*$ désignant le réseau privé du site à l'origine. Maintenant, l'élément de matrice du Hamiltonien apparaissant dans (30.171) est :

$$\langle \chi_1 | H_0 + \Delta V | \chi_{1\vec{R}} \rangle = E_1 \langle \chi_1 | \chi_{1\vec{R}} \rangle + \langle \chi_1 | \Delta V | \chi_{1\vec{R}} \rangle \ , \tag{30.174}$$

d'où la nouvelle forme[44] de (30.170) compte tenu de (30.156) :

$$\varepsilon(\vec{k}) = \frac{\sum_{\vec{R}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \left(E_1 \langle \chi_1 | \chi_{1\vec{R}} \rangle + \langle \chi_1 | \Delta V | \chi_{1\vec{R}} \rangle\right)}{\sum_{\vec{R}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \langle \chi_1 | \chi_{1\vec{R}} \rangle} = E_1 + \frac{\sum_{\vec{R}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \langle \chi_1 | \Delta V | \chi_{1\vec{R}} \rangle}{\sum_{\vec{R}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \langle \chi_1 | \chi_{1\vec{R}} \rangle} \ . \tag{30.175}$$

En raison du caractère localisé des $\chi_{1\vec{R}}(\vec{r})$, les éléments de matrice ont une dépendance grossièrement exponentielle par rapport à la distance $\|\vec{R}\|$ à l'origine. Le moins que l'on puisse faire est donc de retenir les seuls couples correspondant à des sites au plus proches voisins entre eux. On obtient ainsi dans un premier temps :

$$\varepsilon(\vec{k}) = E_1 + \frac{\langle \chi_1 | \Delta V | \chi_1 \rangle + \sum_{\langle \vec{0},\, \vec{R} \rangle} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \langle \chi_1 | \Delta V | \chi_{1\vec{R}} \rangle}{1 + \sum_{\langle \vec{0},\, \vec{R} \rangle} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \langle \chi_1 | \chi_{1\vec{R}} \rangle} \ . \tag{30.176}$$

où les crochets $\langle \vec{0},\, \vec{R} \rangle$ indiquent que la sommation se fait sur les seuls premiers voisins du site origine $\vec{R} = \vec{0}$. Les produits scalaires de la somme du dénominateur sont usuellement petits, et sont négligés le plus souvent. Dans ces conditions, on se retrouve avec :

$$\varepsilon(\vec{k}) = E_1 + \langle \chi_1 | \Delta V | \chi_1 \rangle + \sum_{\langle \vec{0},\, \vec{R} \rangle} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \langle \chi_1 | \Delta V | \chi_{1\vec{R}} \rangle \ ; \tag{30.177}$$

le second terme peut être absorbé dans une redéfinition de E_1, notée ε_0 ; avec ces diverses approximations, le résultat final retenu est :

$$\boxed{\varepsilon(\vec{k}) = \varepsilon_0 + \sum_{\langle \vec{0},\, \vec{R} \rangle} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \langle \chi_1 | \Delta V | \chi_{1\vec{R}} \rangle} \tag{30.178}$$

[44] On relèvera l'analogie entre cette expression et celle des énergies propres du benzène, éq. (28.99).

C'est la dépendance en $\vec{k}$ de la somme qui introduit la dispersion de la bande, laquelle reste faible puisque les éléments de matrice de ΔV restent petits par rapport aux énergies fixant l'ordre de grandeur de ε_0.

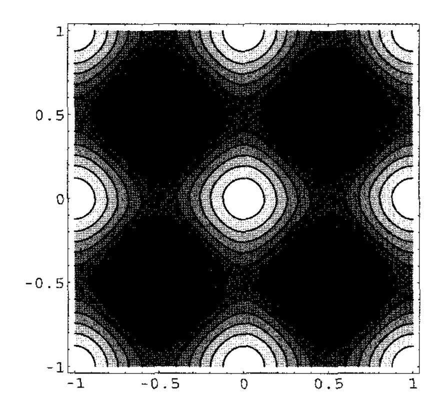


Figure 30.6: Lignes iso-énergie en liaisons fortes pour un réseau carré bi-dimensionnel.

À titre d'exemple, soit un réseau cubique simple dans $\mathbb{R}^3$; le nœud origine a six premiers voisins situés à la distance a et tous les éléments de matrice sont égaux entre eux. En notant $w \in \mathbb{R}$ cette valeur commune, l'expression (30.178) donne :

$$\boxed{\varepsilon(\vec{k}\,) = \varepsilon_0 + 2w\,(\cos k_x a + \cos k_y a + \cos k_z a)} \tag{30.179}$$

Cette expression est visiblement périodique ($\varepsilon(\vec{k} + \vec{K}) = \varepsilon(\vec{k}\,)$). Elle définit une bande d'énergie s'étendant entre les deux valeurs[45] $\varepsilon_0 \pm 6|w|$; mesurée à la bonne aune, la dispersion est effectivement faible (bande étroite).

♦ *Remarques*

1. D'une façon générale, pour tout état de Bloch *exact*[46] $\psi_{\vec{k}}$ et indépendamment de

[45] Le signe de w est sans incidence physique, comme il se doit puisqu'il s'agit d'un élément de matrice non-diagonal. Pour y voir plus clair, il suffit de regarder en détails le cas unidimensionnel, pour lequel l'approximation en cours donne :

$$\psi_k(x) = \frac{1}{\sqrt{N}} \sum_{n\in\mathbb{Z}} \mathrm{e}^{inka} \phi_n(x) \ . \tag{30.180}$$

Le changement de signe de w peut s'obtenir par un rephasage $(-1)^n$ de chaque ϕ_n, ce qui remplace simplement k par $k - \frac{\pi}{a}$ $(\frac{2\pi}{a})$: rien ne change (physiquement), que l'on se place dans la première zone de Brillouin (k dans l'intervalle $[-\frac{\pi}{a}, \frac{\pi}{a}]$) ou dans une zone décalée de $\frac{\pi}{a}$ ($k \in [0, \frac{2\pi}{a}]$).

Très souvent, on choisit $w < 0$ pour faire commodément les petits dessins dans la première zone de Brillouin : on remplit alors la cuvette par le centre, au lieu de remplir l'"anti-cuvette" par les bords...

[46] c'est-à-dire obtenu en résolvant exactement l'équation pour $u_{\vec{k}}$.

toute approximation, on peut définir sa transformée de Fourier par :

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} \phi_{\vec{R}}^{(\mathrm{W})}(\vec{r}) \Longleftrightarrow \phi_{\vec{R}}^{(\mathrm{W})}(\vec{r}) \stackrel{\text{déf}}{=} \frac{1}{\sqrt{N}} \sum_{\vec{k}\in BZ1} \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}} \psi_{\vec{k}}(\vec{r}) \ . \tag{30.181}$$

La fonction $\phi_{\vec{R}}^{(\mathrm{W})}$ ainsi définie porte le nom de fonction de Wannier. Contrairement aux fonctions atomiques $\chi_{\nu\vec{R}}$ ou $\phi_{\vec{R}}$ introduites ci-dessus, deux fonctions de Wannier relatives à deux sites distincts sont orthogonales :

$$\langle\phi_{\vec{R}}^{(\mathrm{W})}|\phi_{\vec{R}'}^{(\mathrm{W})}\rangle = \delta_{\vec{R}\vec{R}'} \ , \tag{30.182}$$

et constituent une autre base possible de représentation ; l'orthogonalité des fonctions de Wannier résulte du fait que la transformation (30.181) est unitaire, les $\psi_{\vec{k}}(\vec{r})$ étant mutuellement orthogonales. En tant que transformée de Fourier d'une fonction étendue, $\phi_{\vec{R}}^{(\mathrm{W})}(\vec{r})$ est une fonction présentant une forte localisation au voisinage de $\vec{R}$, un caractère qu'elle partage avec les orbitales atomiques $\chi_{\nu\vec{R}}$ ou les fonctions $\phi_{\vec{R}}$ introduites en (30.146).

2. On aura réalisé que la méthode utilisée est typiquement du genre LCAO : on fabrique des *orbitales moléculaires* pour un électron dans le solide, à partir de fonctions localisées du genre orbitale atomique. Au fond, le schéma ici utilisé est exactement de même nature que celui développé pour le benzène (chapitre 28, sous-section 28.2.4) et pour le polyacétylène (ibid., sous-section 28.2.5). Cette description est donc fortement contrastée vis-à-vis de la méthode de Heitler - London (ibid., sous-section 28.2.3) ; ici, cette méthode consisterait à considérer d'abord les *produits* $\prod_{\vec{R}} \phi_{\vec{R}}(\vec{r}_i)$ avant de diagonaliser tant bien que mal et enfin former les bonnes combinaisons linéaires vis-à-vis de la symétrie de permutation. On voit de suite que les fonctions ainsi construites contiendraient des configurations atomiques d'excitation arbitraire, ce qui ne serait pas convenable physiquement pour décrire approximativement un réseau où, par hypothèse, les éléments non-diagonaux du Hamiltonien, mesurant l'importance des sauts d'un site à l'autre, sont très petites devant une excitation atomique. ♦

30.5 Conducteurs, semi-conducteurs, isolants

Dans cette section, il s'agit principalement de comprendre l'origine de la conduction électrique des solides, mesurée par la conductivité σ. Celle-ci varie énormément d'un matériau à l'autre : il existe un facteur 10^9 entre la conductivité du cuivre et celle du germanium pur à l'ambiante, et un facteur 10^{19} entre celle du cuivre et celle du verre ordinaire. Il est rare qu'une même propriété physique varie dans un intervalle aussi gigantesque et c'est ce fait qui légitime la définition de trois classes de matériaux : les conducteurs, les semi-conducteurs et les isolants[47]. Cette différenciation est expliquée

[47] Ces énormes différences d'ordre de grandeur traduisent les différences profondes de la nature du transport de l'électricité dans ces divers solides. Pour un métal, le transport est possible grâce à la

par la théorie quantique, et elle seulement : elle ne peut se décrire simplement en termes de propriétés élémentaires dans l'espace réel et impose de se placer dans l'espace des impulsions (espace réciproque), plus précisément d'examiner la dépendance de l'énergie d'un électron en fonction de sa (pseudo-) impulsion. En outre, elle résulte d'un jeu subtil entre ces lois de dispersion et la statistique quantique, étant entendu que les porteurs de charge (dans le cas du transport ordinaire) sont des électrons, c'est-à-dire des fermions. La structure géométrique du solide joue un rôle important en imposant une loi de dispersion ou une autre, mais la nature physique (structure électronique) des objets remplissant la structure a également un rôle déterminant. Globalement, la théorie des bandes fournit le cadre théorique permettant de comprendre pourquoi, quand les électrons interagissent peu, un matériau est conducteur ou non.

Les deux situations extrêmes qui viennent d'être étudiées mettent en évidence un phénomène indissolublement lié à l'existence du potentiel périodique : l'ouverture de petits *gaps* en énergie par comparaison avec la distribution continue en énergie, de 0 à $+\infty$, pour un électron libre faiblement perturbé par le potentiel périodique, ou l'apparition d'une (faible) dispersion des états atomiques pour un électron fortement lié. Dans tous les cas, en raison exclusive de la périodicité spatiale et du théorème de Bloch qui en résulte, le parcours de l'axe des énergies fait rencontrer des zones où il existe des états (bandes *permises*) et d'autres où il n'y en a pas (bandes *interdites*).

Cette propriété peut être utilement précisée à l'aide de la notion de *densité d'états*, $\mathcal{D}(E)$; par définition :

$$\boxed{\mathcal{D}(E_0)\,\delta E \stackrel{\text{déf}}{=} \text{Nombre d'états d'énergie comprise entre } E_0 \text{ et } E_0 + \delta E} \tag{30.183}$$

$\mathcal{D}(E_0)\,\delta E$ est une mesure de la dégénérescence et donne le nombre d'états de même énergie E_0, à δE près, se distinguant les uns des autres par les autres nombres quantiques nécessaires pour spécifier complètement l'état quantique. Un électron dans un potentiel périodique est caractérisé par sa pseudo-impulsion $\hbar\vec{k}$, et son spin. En l'absence de champ magnétique, ou de tout effet spin-orbite, l'énergie d'un électron dans une bande donnée, $\varepsilon_n(\vec{k})$, ne dépend pas du spin. Le nombre d'états d'énergie E_0 à δE près s'obtient en sommant sur toutes les valeurs de $\vec{k}$ donnant $\varepsilon_n(\vec{k}) = E_0$, et sur le spin. Ce dernier fournit un simple facteur 2, d'où ($\delta^3 k = (\frac{2\pi}{L})^3$) :

$$\boxed{\sum_{\text{états}} \equiv 2\sum_{\vec{k}} \quad \rightarrow \quad 2\left(\frac{L}{2\pi}\right)^3 \int \mathrm{d}^3 k} \tag{30.184}$$

présence d'électrons presque libres, et la conductivité *décroît* avec la température : plus la température est élevée, plus les ions du métal vibrent avec une grande amplitude et plus les électrons sont freinés. Pour un semi-conducteur intrinsèque – dont la résistivité *décroît* avec la température –, les porteurs de charge sont encore des électrons, mais seule l'activation thermique permet à une (petite) fraction d'entre eux de franchir les (petits) *gaps* d'énergie entre la bande de valence et la bande de conduction. Enfin, dans le cas d'un verre, ce sont des ions qui transportent le courant ; la diffusion d'espèces très massives permet à elle seule de se douter que la résistivité est extrêmement élevée – toutes choses égales par ailleurs.

En conséquence, la densité d'états de la bande d'énergie $\varepsilon_n(\vec{k}\,)$ est donnée par[48] :

$$\mathcal{D}_n(E_0) = 2L^3 \int_{\mathbb{R}^3} \frac{\mathrm{d}^3k}{(2\pi)^3}\, \delta(\varepsilon_n(\vec{k}) - E_0) \qquad ([\mathcal{D}] = \mathrm{E}^{-1}) \ . \tag{30.185}$$

Autrement dit, $\mathcal{D}_n(E)$ résulte de la sommation sur tous les états possibles filtrée à l'aide de la fonction de Dirac. On sait que pour une fonction monotone[49] :

$$\delta(f(x) - f_0) = \frac{1}{|f'(x_i)|}\, \delta(x - x_i) \ ; \tag{30.187}$$

à trois dimensions, cette relation se généralise en[50] :

$$\delta\big(f(\vec{x}) - f_0\big) = \frac{1}{\|\vec{\nabla} f(\vec{x}_0)\|}\, \delta\big((\vec{x} - \vec{x}_0).\vec{n}\big) \ , \qquad f(\vec{x}_0) = f_0 \ , \tag{30.191}$$

où $\vec{n}$ est le vecteur unitaire de la normale à la surface S_{f_0} d'équation $f(\vec{x}) = f_0$. Avec ces résultats, la densité (30.185) s'écrit :

$$\mathcal{D}_n(E_0) = L^3 \int_{\mathbb{R}^3} \frac{\mathrm{d}^3k}{4\pi^3}\, \frac{1}{\|\vec{\nabla}\varepsilon_n(\vec{k}_0)\|}\, \delta((\vec{k} - \vec{k}_0).\vec{n}) \ , \qquad \varepsilon_n(\vec{k}_0) = E_0 \ . \tag{30.192}$$

La présence de la fonction de Dirac dit que seuls les points où $(\vec{k} - \vec{k}_0).\vec{n}$ est nul contribuent à l'intégrale, c'est-à-dire tous les points extrémités de $\vec{k}$ dont l'écart à $\vec{k}_0$ compté le long de la normale s'annule : il s'agit donc des points du plan tangent infiniment proches de $\vec{k}_0$. Quand $\vec{k}_0$ varie, l'ensemble de ces points balaie la surface d'énergie E_0. Par suite, une expression équivalente à (30.192) pour la densité d'états $\mathcal{D}_n(E)$ est :

$$\boxed{\mathcal{D}_n(E) = L^3 \int_{S_E} \frac{\mathrm{d}^2k}{4\pi^3}\, \frac{1}{\|\vec{\nabla}\varepsilon_n(\vec{k}\,)\|}} \tag{30.193}$$

[48]La densité d'états est parfois définie par unité de volume, ce qui revient à poser $L = 1$ dans la définition (30.185).

[49]Si la fonction n'est pas monotone, l'égalité $f(x) = f_0$ peut être satisfaite pour plusieurs valeurs x_i de x ; alors :

$$\delta(f(x) - f_0) = \sum_i \frac{1}{|f'(x_i)|}\, \delta(x - x_i) \ , \qquad f(x_i) = f_0 \quad \forall i \ . \tag{30.186}$$

[50]En effet, le développement de Taylor au voisinage de $\vec{x}_0 \in S_{f_0}$ s'écrit :

$$f(\vec{x}) = f_0 + (\vec{x} - \vec{x}_0).\vec{\nabla} f(\vec{x}_0) + \mathcal{O}\big((\vec{x} - \vec{x}_0)^2\big) \ . \tag{30.188}$$

Par définition, $\vec{\nabla} f$ est perpendiculaire à la surface S_{f_0}, donc dirigé suivant la normale $\vec{n}$; il en résulte que (30.188) se réécrit comme :

$$f(\vec{x}) = f_0 + \|\vec{\nabla} f(\vec{x}_0)\|(\vec{x} - \vec{x}_0).\vec{n} + \mathcal{O}\big((\vec{x} - \vec{x}_0)^2\big) \ , \tag{30.189}$$

ce qui donne :

$$\delta\big(f(\vec{x}) - f_0\big) = \delta\big(\|\vec{\nabla} f(\vec{x}_0)\|(\vec{x} - \vec{x}_0).\vec{n}\big) \ , \tag{30.190}$$

et compte tenu de (30.187), on trouve (30.191). On vérifie sans peine l'homogénéité de (30.191), sachant que $[\delta(\vec{x}.\vec{n})] = \mathrm{L}^{-1}$ (alors que $[\delta(\vec{x} \in \mathbb{R}^3)] = \mathrm{L}^{-3}$).

où S_E est la surface d'équation $\varepsilon_n(\vec{k}) = E$. Ces définitions se généralisent en toute dimension. Par exemple, pour un réseau plan, la densité $\mathcal{D}$ s'obtient en effectuant une intégrale de *ligne* et non de surface, le long de la ligne d'énergie constante $\varepsilon_n(\vec{k}) = E$, $\vec{k}$ étant maintenant un vecteur à deux composantes k_x, k_y.

À l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$, $\varepsilon_n(\vec{k})$ présente au moins un minimum et un maximum, en tant que fonction continue différentiable et périodique. En ces *extrema*, $\vec{\nabla}\varepsilon_n$ s'annule et l'intégrand diverge. À $d = 3$, ces singularités sont intégrables et $\mathcal{D}_n(E)$ reste une fonction continue, tout en présentant des points anguleux appelés singularités de van Hove. Au contraire, à $d = 1$, quand $\vec{\nabla}_{\vec{k}}\varepsilon_n(k)$ s'annule linéairement, $\mathcal{D}_n$ diverge logarithmiquement.

Pour une énergie donnée E, la densité totale s'obtient en faisant simplement la somme des densités des bandes :

$$\boxed{\mathcal{D}(E) = \sum_n \mathcal{D}_n(E)} \tag{30.194}$$

de sorte que l'allure générale de $\mathcal{D}(\mathcal{E})$ est celle indiquée sur la figure 30.7, sans le souci de faire figurer les détails (et en particulier les singularités de van Hove). Bien sûr, il peut arriver – c'est fréquent – que deux bandes se recouvrent en énergie (voir fig. 30.7, à droite). En tout état de cause, le spectre des énergies à une particule d'un solide périodique n'est ni simplement connexe, ni purement ponctuel. En quelque sorte, c'est un hybride entre celui d'un électron libre et celui d'un électron lié dans un atome.

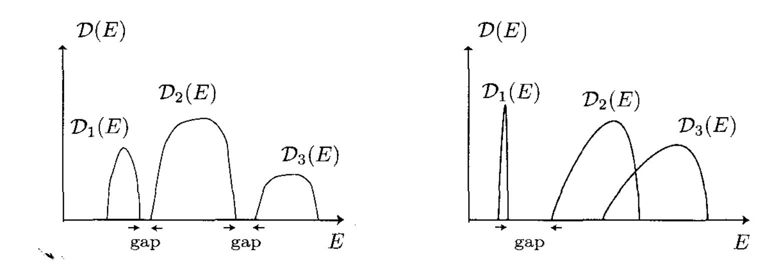


Figure 30.7: Illustration schématique de la densité d'états. À droite, l'exemple de deux bandes qui se recouvrent.

La densité d'états $\mathcal{D}(E)$ est tout particulièrement utile pour calculer les valeurs moyennes de grandeurs ne dépendant *que* de l'énergie ; en effet, $A(E)$ désignant une telle grandeur, sa moyenne contient des sommes du genre :

$$\langle A \rangle = \sum_{\text{états}} A(E) \quad \rightarrow \quad 2\Big(\frac{L}{2\pi}\Big)^3 \int \mathrm{d}^3k\, A(E) \ , \tag{30.195}$$

où le facteur 2 prend en compte le spin (ce qui suppose en outre que $A(E)$ ne dépend pas non plus du spin). L'intégration peut s'effectuer en deux temps : pour une bande

d'énergie de loi de dispersion $\varepsilon(\vec{k})$, on fixe la valeur de l'énergie et on intègre sur $\vec{k}$ à E fixée :

$$\langle A\rangle = 2\Big(\frac{L}{2\pi}\Big)^3 \int \mathrm{d}E \int \delta\big(\varepsilon(\vec{k}) - E\big)\, A(E)\, \mathrm{d}^3k \ . \tag{30.196}$$

Avec les règles habituelles concernant la fonction de Dirac rappelées plus haut, (30.196) devient :

$$2\Big(\frac{L}{2\pi}\Big)^3 \int \mathrm{d}E \int_{S_E} \frac{1}{\|\vec{\nabla}\varepsilon(\vec{k})\|} A(E)\, \mathrm{d}^2k \ , \tag{30.197}$$

de sorte que :

$$\langle A\rangle = \int \mathcal{D}(E)\, A(E)\mathrm{d}E \equiv \int \tilde{A}(E)\mathrm{d}E \ . \tag{30.198}$$

$\tilde{A}(E)$ étant une observable effective incorporant la densité d'états sommée sur les degrés de liberté dont $A(E)$ ne dépend pas.

Les distinctions quantitatives des deux situations extrêmes étudiées plus haut peuvent s'illustrer comme indiqué sur la figure 30.8 : pour les électrons presque libres, les bandes sont larges et séparées par des petits *gaps* ; la situation est inverse pour un électron en liaisons fortes.

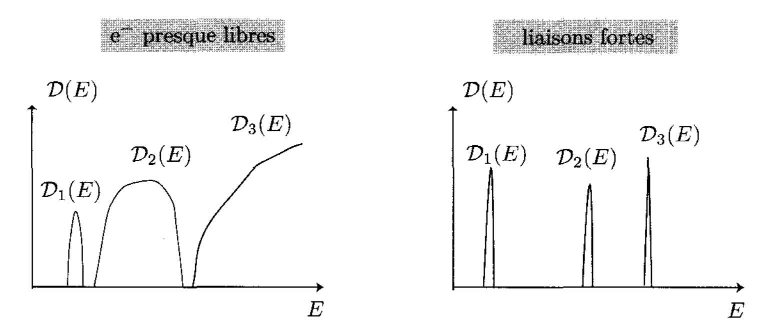


Figure 30.8: Illustration schématique des différences de la densité d'états pour les approximations des électrons presque libres (à gauche) et des liaisons fortes (à droite).

Enfin, une fois muni de la densité totale d'états $\mathcal{D}(E)$, (30.194), il est possible d'écrire formellement l'équation donnant l'énergie de Fermi ε_F. Si N_e est le nombre d'électrons à caser, le niveau de Fermi est déterminé (à température nulle) par l'égalité :

$$\boxed{\int_0^{\varepsilon_\mathrm{F}} \mathcal{D}(E)\, \mathrm{d}E = N_\mathrm{e}} \tag{30.199}$$

On est maintenant en mesure de comprendre – en raisonnant essentiellement à température nulle[51] – l'existence de deux grandes classes de matériaux, les conducteurs

[51]Pour les métaux usuels, ε_F est de l'ordre de 10^4 eV ou plus, et on a toujours $k_\mathrm{B}T \ll \varepsilon_\mathrm{F}$; la distribution de Fermi reste donc toujours très voisine de la fonction de Heaviside $Y(\varepsilon_\mathrm{F} - E)$.

et les isolants. La discussion repose crucialement sur l'hypothèse de la validité d'une théorie à une particule (ou d'une théorie de champ moyen), ce qui permet de s'appuyer sur l'existence d'états propres bien définis pour un seul électron (nu ou habillé), mais essentiellement corrélé aux autres par le Principe de Pauli.

Il s'agit de construire l'état fondamental du cristal, étant donné un certain nombre d'électrons à caser, N_e. Ces derniers proviennent usuellement de la couche de valence des atomes qui s'assemblent pour former le cristal. Chaque atome fournit un petit nombre d'électrons susceptibles de participer au transport électrique, n_e ($n_e = 1$ pour un alcalin, $n_e = 8$ pour Fe, dont la configuration électronique est [Ar] $(3d)^6(4s)^2$) ; s'il y a N mailles primitives et un atome par maille (cristal monoatomique), il y a donc $N_e = n_e N$ électrons à placer dans les bandes à une particule, $\varepsilon_n(\vec{k})$, chacune d'entre elles pouvant accueillir $2N$ électrons, compte tenu des deux possibilités de spin.

Construire l'état fondamental, c'est donc remplir les états de Bloch les uns après les autres, en se déplaçant évidemment dans le sens des énergies croissantes, et en piochant dans le *stock* d'électrons disponibles jusqu'à épuisement. Pour la clarté des choses, on supposera dans un premier temps que les bandes ne se recouvrent pas en énergie. Avec cette hypothèse, il existe une et une seule plus haute bande en énergie contenant des électrons, appelée *bande de conduction*[52]. Deux cas sont alors possibles :

1. cette plus haute bande est pleine et contient donc $2N$ électrons. Alors, tout électron est coincé en énergie, sauf à gagner une énergie au moins égale à celle du *gap*, E_g, qu'aucun champ extérieur ne peut lui fournir[53]. À température nulle, le solide est donc un isolant. À température finie, les fluctuations thermiques ne peuvent rien changer tant que $k_B T \ll E_g$; comme $(k_B T)_{\text{ambiante}} \simeq 25$ meV, le changement de comportement conductif induit par la température finie dépend de la valeur numérique de E_g :

    - si E_g est de l'ordre de quelques eV (pour le diamant $E_g \simeq 5,5$ eV), les probabilités d'excitation varient en gros comme $e^{-40\times 5,5}$: autant dire qu'elles sont nulles. Le solide est isolant à toute température où il existe en tant que tel[54].
    - Si E_g est nettement plus petit, une fraction d'eV pour fixer les idées (exemple : le germanium pour lequel $E_g \simeq 0,67$ eV à l'ambiante[55]), les probabilités sont de l'ordre de e^{-30}, ce qui est encore très petit. Cependant, compte tenu de la relativement haute densité de porteurs potentiels, la conductivité devient notable, quoique très petite par rapport à un bon conducteur comme le

[52] l'avant-dernière bande complètement remplie s'appelle *bande de valence*.

[53] Cette affimation sera validée ultérieurement, mais on sait bien (voir l'effet Stark, chapitre 23, sous-section 23.3.1, notamment éq. (23.171)) que les champs électriques usuels du laboratoire sont microscopiques par rapport à un champ intra-atomique ; ceci reste vrai même quand on passe de l'atome au solide.

[54] En pratique, l'affirmation doit évidemment être nuancée, notamment parce que le solide peut contenir des impuretés faisant apparaître des états dans les *gaps*.

[55] La dilatation du réseau avec la température T donne une dépendance en T aux énergies des *gaps*.

cuivre. Le solide est alors dit *semi-conducteur* (intrinsèque[56]). La résistivité d'un semi-conducteur *décroît* quand la température augmente, comportement inverse de celui observé dans les métaux.

2. La plus haute bande est incomplètement remplie, ce qui situe le niveau de Fermi ε_F à l'intérieur de cette bande à la hauteur du dernier niveau occupé (niveau d'eau de la mer de Fermi). Les électrons qui se trouvent dans le voisinage inférieur de ε_F se voient offerts des niveaux d'énergie infiniment proches ; dès lors, un champ électrique, aussi petit soit-il, peut de fait les mettre en mouvement et leur donner une énergie *orientée* – qui reste toutefois minuscule par rapport à l'énergie thermique – donnant lieu à un vecteur courant, donc au passage d'un courant électrique : le solide est un métal.

À l'intérieur d'une maille primitive de $\tilde{\mathcal{B}}$, il y a autant d'états $\vec{k}$ distincts qu'il y a de mailles primitives dans $\mathcal{B}$, et chacun peut recevoir deux électrons compte tenu du spin. Si $\mathcal{B}$ n'a pas de base (cristal mono-atomique) et si chaque atome fournit un seul électron ($n_e = 1$), alors la dernière bande est à moitié remplie (*demi-remplissage*) – toujours dans l'hypothèse de bandes qui ne se recouvrent pas – et le cristal est conducteur ; c'est encore vrai si n_e est impair. La situation est inversée si n_e est pair, et le solide est isolant.

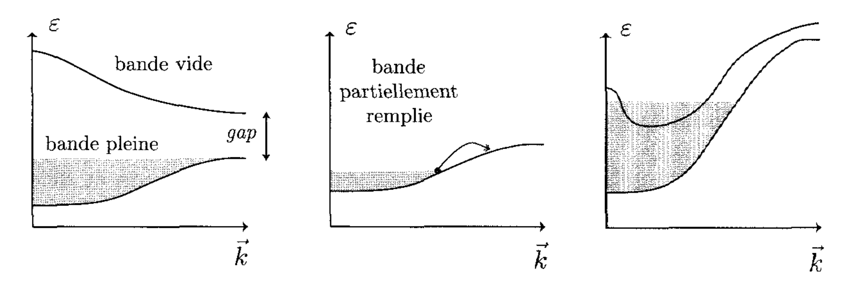


Figure 30.9: Illustration schématique selon la théorie des bandes (théorie à *une* particule) de la différence entre isolant (à gauche), métal (au milieu) et semi-métal (à droite).

L'existence de recouvrement[57] entre les bandes modifie sérieusement ces dernières conclusions. En pareil cas, si n_e est impair, les dernières bandes se partagent les électrons et on n'obtient évidemment pas un demi-remplissage ; toutefois, rien n'est changé fondamentalement en ce qui concerne les propriétés de conduction : le solide reste conducteur. Au contraire, quand n_e est pair, un solide qui est isolant en l'absence de recouvrement

[56]Cette précision est là pour distinguer le semi-conducteur d'un matériau où l'on a délibérément introduit des impuretés (dopage, semi-conducteur *extrinsèque*). Cette opération produit des états à l'intérieur du *gap*, qui sont autant de barreaux d'échelle permettant le franchissement par étapes de ce dernier.

[57]Un cas limite est celui où deux bandes, l'une pleine, l'autre vide, entrent en contact sans toutefois se dépasser l'une l'autre en un nombre fini de points de BZ1. En pareil cas, le solide n'est pas conducteur (toujours à température nulle) puisque le canal dans l'espace de $\vec{k}$ permettant de passer de la bande pleine à la bande vide est "infiniment" étroit (dit autrement, la frontière à l'énergie de Fermi ε_F est de mesure nulle).

devient conducteur puisque les deux bandes qui se recouvrent sont inévitablement maintenant incomplètes[58] (voir fig. 30.9, à droite).

Il importe de réaliser le rôle crucial joué par la *dispersion* des bandes électroniques, c'est-à-dire le fait que leur énergie ε dépend de $\vec{k}$: un électron dans une bande plate ne peut pas gagner d'énergie sous l'effet d'un champ, même si la bande est quasi-vide. Quand la bande est partiellement remplie, les électrons peuvent certes passer d'un état $\vec{k}$ à un état $\vec{k}\,'$ sous l'effet de collisions, mais leur énergie ne change pas puisqu'elle est précisément indépendante de la pseudo-impulsion. La "platitude" de l'énergie d'une bande se reporte immédiatement sur la vitesse (moyenne) d'un électron de Bloch, en vertu de la relation $\vec{v} = \hbar^{-1}\vec{\nabla}_{\vec{k}}\varepsilon$: quand la bande est plate, la vitesse est nulle.

Dans les cas moins extrêmes, les grandeurs relatives des composantes de $\vec{\nabla}_{\vec{k}}\varepsilon$ sont l'une des causes premières des anisotropies de conductivité (d'autres facteurs doivent être pris en compte, notamment la densité d'états au niveau de Fermi). Pour un solide qui se clive facilement (le carbone graphite, par exemple), cette propriété de clivage provient principalement du fait que la distance entre deux plans du réseau en nid d'abeille est plusieurs fois supérieure[59] à la plus petite distance entre deux atomes de carbone d'un même plan. Comme la dispersion des bandes est pilotée par les amplitudes de saut – lesquelles ont en gros une dépendance exponentielle en distance –, il en résulte que la dispersion des bandes dans la direction perpendiculaire aux plans est beaucoup plus faible que parallèlement aux plans, d'où une forte anisotropie de conductivité. Notons enfin qu'une bande plate donne *ipso facto* une masse effective infinie, une autre façon de garder à l'esprit l'importance fondamentale de la dispersion.

On retiendra également le lien semi-quantitatif entre dispersion et délocalisation des états. Pour les électrons presque libres, la dispersion est grande, les états étant des ondes planes modulées en amplitude par la fonction $u_{\vec{k}}$; il en résulte que la densité électronique est très *diffuse*[60] et que, si on prépare un paquet d'ondes localisé (sur un seul site, par exemple), il s'étale très vite. Au contraire, pour des électrons en liaisons fortes, la dispersion est faible ; les états propres sont toujours étendus, puisqu'il s'agit d'états de Boch, mais ils présentent une forte structure spatiale, traduisant le fait que l'électron, en quelque sorte, passe par effet-tunnel d'un site à l'autre. Ici encore, un paquet d'ondes localisé finit par se diluer sur le réseau, mais l'évolution se fait sur une échelle de temps beaucoup plus longue ($\sim \hbar/t$, t désignant l'amplitude de saut élémentaire). À la limite d'une amplitude de saut infiniment petite, l'énergie ne dépend pas de $\vec{k}$, la dégénérescence est $\sim N$ et les fonctions atomiques localisées sont, tout autant que les états de Bloch ou n'importe quelle autre combinaison linéaire, des fonctions propres du Hamiltonien d'un électron. En définitive, la dispersion ou la non-dispersion des bandes est un indice commode de la nature, localisée ou étendue, des états électroniques.

[58] Par exemple, pour un cristal carré à deux dimensions contenant deux électrons par maille, la sphère (cercle) de Fermi déborde de BZ1, de sorte que deux bandes se trouvent partiellement remplies et le solide est conducteur.

[59] environ 2,4 fois plus grande. Comme les intégrales de saut ont une variation grossièrement exponentielle avec la distance, ceci donne une réduction d'un facteur ~ 10 pour les amplitudes de saut d'un plan à l'autre.

[60] On parle alors usuellement d'états *étendus*.

♦ *Remarques*

1. En toute généralité, le niveau de Fermi se définit comme la limite du potentiel chimique à température nulle :

$$\boxed{\varepsilon_{\rm F} \stackrel{\rm déf}{=} \lim_{T\rightarrow 0} \mu(T, V, N)} \tag{30.200}$$

   Pour un métal, $\varepsilon_{\rm F}$ ainsi défini est, effectivement et sans ambiguïté, l'énergie du plus haut état occupé. Dans le cas d'un semi-conducteur ou d'un isolant, la limite $\mu(T \rightarrow 0_+, V, N)$ est située quelque part *dans* le *gap*, entre le sommet de la bande de valence et le bas de la bande de conduction, à un certain endroit qui dépend de leur symétrie (dégénérescence) et de la masse effective (concavité), c'est-à-dire finalement de la densité d'états $\mathcal{D}(E)$. Si celle-ci est la même dans les deux bandes, $\varepsilon_{\rm F} = \mu(T \rightarrow 0_+, V, N)$ est au milieu du *gap*. Formellement, ceci permet de lever l'ambiguïté sur la valeur à retenir physiquement pour la marche de Fermi $\theta(\varepsilon_{\rm F} - \varepsilon)$ en $\varepsilon = \varepsilon_{\rm F}$.

2. Il est souvent commode de considérer l'état fondamental comme un état "vide" (d'excitations). En ce sens, la mer de Fermi est la cousine de la mer de Dirac – que l'on a supposée complètement remplie pour limiter les ennuis avec les états d'énergie négative. À partir de cet état vide, l'excitation de basse énergie (par des photons, par exemple) crée des paires électron-trou, analogues (sur une autre échelle d'énergie !) des paires électron-positron de la théorie de Dirac. Pour un isolant ou un semi-conducteur, ces transitions bande à bande se produisent pour une fréquence du champ électromagnétique appliqué supérieure ou égale à $h^{-1}\Delta E_{\rm g}$, appelé *seuil de photoconductivité*[61]. La mesure de cette fréquence-seuil est un moyen courant pour obtenir la valeur numérique du *gap*.

3. Comme rappelé au début de la discussion, cette explication élémentaire de la distinction entre conducteurs et isolants repose crucialement sur l'hypothèse de la validité d'une théorie à une particule, où les interactions entre électrons sont telles que la notion de quasi-particule est préservée. Lorsque tel n'est pas le cas, la question de la distinction isolant vs. conducteur doit être entièrement reconsidérée. Il est facile de comprendre pourquoi en raisonnant physiquement par référence à un cas-limite.

   Pour simplifier la discussion, supposons que chaque atome apporte un seul électron (demi-remplissage), auquel cas, selon la théorie des bandes, on est en présence d'un conducteur. Introduisons maintenant l'interaction entre les électrons (on sort du schéma des bandes) et admettons que leur répulsion mutuelle (traditionnellement notée U) est si forte que leur énergie cinétique peut être négligée devant l'énergie potentielle d'interaction. Il est bien clair, alors, que chaque électron a envie de rester sur un atome : dès qu'il veut en bouger, les autres le repoussent fortement.

[61] Il s'agit bien de *photoconductivité* : l'exposition à la lumière – de fréquence suffisante, encore une preuve expérimentale du photon ! – porte des électrons dans la bande de conduction initialement vide : l'isolant, ou le semiconducteur, devient un vrai conducteur, la conductivité augmentant avec le flux lumineux, toutes choses égales par ailleurs.

En définitive, dans cette image extrême, les électrons eux-mêmes ont tendance à former un réseau statique, coïncidant pour cette situation de demi-remplissage, au réseau de Bravais sous-jacent du solide[62]. Il est physiquement évident qu'un tel solide ne peut être conducteur – tous les porteurs de charge sont bloqués. Il est non moins évident qu'il n'y a nulle contradiction : le schéma explicatif en terme de bande est invalide, puisqu'il résulte d'une théorie à *une* particule. Pour cette raison, il s'agit alors d'une classe particulière d'isolant, appelé *isolant de Mott.*

Dans un cas moins extrême, la discussion repose sur les valeurs comparées de la répulsion électronique – d'ordre U quand deux électrons sont sur le même site – et de l'énergie cinétique, dont l'ordre de grandeur est noté (traditionnellement) t. Tout (ou presque) dépend du rapport $\alpha \stackrel{\text{déf}}{=} t/U$. Conceptuellement, on peut imaginer une variation graduelle de α. Quand $\alpha \gg 1$, le solide considéré ci-dessus est sûrement conducteur, la théorie des bandes retrouvant sa validité dans cette limite ; quand α est petit, on a un isolant, d'après l'argument qualitatif ci-dessus. Le passage d'une phase à l'autre porte le nom de transition de Mott, où l'on voit la conductivité chuter brutalement à zéro quand α décroît au voisinage d'une certaine valeur critique α_{C}, comportement réminiscent d'une *transition de phase.*

Physiquement, on peut imaginer que la diminution graduelle de α se produit en dilatant progressivement le réseau d'un alcalin, puisque, typiquement, t décroît exponentiellement avec la distance. Pour les valeurs usuelles de la maille, le solide est conducteur ; distendre le réseau fait diminuer α. Quand le paramètre a de la maille est très grand, les atomes se trouvent très loin les uns des autres et la structure est celle d'un antiferromagnétique[63], les atomes portant chacun un électron dont le spin alterne d'un site à l'autre (ceci pour l'état singulet de spin total nul). À l'inverse, quand a décroît, t augmente puisque le passage d'un site à l'autre est facilité, le rapport t/U croît et la transition de Mott peut se produire, le cas échéant. En pratique, la variation du paramètre de réseau a peut s'obtenir en comprimant le cristal : sous très haute pression, un isolant de Mott peut basculer dans l'état conducteur.

4. Dans un tout autre ordre d'idées, l'explication de la distinction isolant - conducteur donnée plus haut repose également sur une autre hypothèse importante, celle

[62]Dans le régime des faibles densités ($r_{\text{s}} \gg a_0$), c'est le réseau de Wigner.

Quand on n'est pas à demi-remplissage, la densité d'électrons n'est pas commensurable avec la densité d'atomes, et la physique devient beaucoup plus riche (possibilité d'ondes de densité de charge, appelées CDW, acronyme de *charge density waves*). Ce problème ressemble à une chaîne de boules et de ressorts soumis en plus à un potentiel périodique dont la maille élémentaire n'est pas commensurable avec le réseau formé par les positions d'équilibre des boules de la chaîne (modèle de Frenkel - Kontorova). Il est aisé d'appréhender intuitivement l'extrême complexité d'un tel système (pour une animation, voir [246]).

[63]Cette affirmation pour le moins rapide laisse de côté de grandes difficultés. En effet, la structure antiferromagnétique (brièvement rediscutée plus loin, voir section 30.6) ne s'impose pas en l'état actuel des choses puisque tout basculement de spins conservant $M_S = 0$ ne coûte rien ; on peut même imaginer une structure en domaines (des groupes de spins tous alignés, soit vers le haut, soit vers le bas) : ce que fait le système dépend crucialement de l'interaction effective entre spins, qui donne toute sa richesse au problème. En définitive, dans la limite $\alpha = 0$, la dégénérescence est infinie. Savoir ce que fait réellement le système à α petit est une question d'une grande difficulté. Le fondamental du Hamiltonien de Hubbard a été obtenu en 1968 par Lieb et Wu [241], [242] par la méthode de l'*Ansatz* de Bethe ; ces auteurs ont montré que dans le strict cadre du modèle de Hubbard, il n'y a pas de transition de Mott.

d'un réseau *rigide*. Dans certains cas, autoriser une déformation de réseau peut donner lieu à une modification spectaculaire de l'image fournie par la théorie des bandes avec un réseau fixé. L'exemple le plus connu est l'instabilité dite de Peierls, dont on a évoqué les fondements à propos du polyacétylène (chapitre 28, sous-section 28.2.5).

L'idée est facile à comprendre dans le cas très simple d'un réseau unidimensionnel à un électron par maille. Si le réseau est rigide, la première zone de Brillouin s'étend entre $\pm\frac{\pi}{a}$ et est à moitié remplie : le solide est conducteur. Imaginons maintenant que l'atome initialement en na est déplacé d'une petite quantité $(-1)^n\delta a$. Autrement dit, on *dimérise* le réseau, les atomes $2n$ et $2n+1$ étant maintenant distants de $a - 2\delta a$. Dès lors, le réseau de Bravais a une maille élémentaire de dimension $2a$ et la nouvelle première zone de Brillouin s'étend entre $\pm\frac{\pi}{2a}$. En repliant la bande d'énergie considérée[64], on voit qu'il y a une stricte dégénérescence en bord de zone ($k = \pm\frac{\pi}{2a}$) ; c'est donc à ces endroits que les *gaps* vont s'ouvrir quand le réseau se déforme. Maintenant, la bande inférieure est pleine, la bande supérieure est vide, un *gap* existe et le solide est isolant !

Il s'agit bien d'une instabilité au sens où l'énergie du gaz d'électrons a spontanément tendance à décroître dès que la dimérisation est rendue possible[65]. Cet abaissement d'énergie entre en compétition avec l'énergie élastique de déformation de réseau, puisque déformer le réseau coûte de l'énergie. En faisant le bilan, on trouve que l'énergie totale présente un minimum pour une certaine valeur δa_0, qui fixe le taux de dimérisation du réseau, la conclusion restant inchangée : dans ce cas précis, autoriser le réseau à se déformer transforme un conducteur en isolant (voir problèmes 31.9.4 et 28.4.5). ♦

30.6 Les magnons : un exemple de l'ordre magnétique d'un solide

L'ordre inhérent à la phase cristalline a des conséquences d'une extrême importance, dont on vient de voir l'une des toutes premières, à propos de la conductibilité électrique d'un solide ; en marge de ce problème ont été notamment évoquées les ondes de densité de spin (acronyme SDW, pour *spin-density wave*). De fait, dès que l'on considère les attributs magnétiques des entités microscopiques constituant le solide, des structures très variées sont susceptibles de se former. Cette section est consacrée à l'une d'entre elles, les *magnons* dans un système réputé ferromagnétique, dont la nature et les propriétés s'inscrivent tout naturellement dans le formalisme développé dans le présent chapitre.

En laissant de côté le diamagnétisme universel (mais donnant lieu à des effets très petits – voir Tome I, chapitre 11, notamment p. 369), on sait à un niveau élémentaire

[64]Après tout, rien n'interdit de représenter le réseau *non déformé* comme un réseau de maille double avec deux atomes par maille. On retrouvera cette façon de faire lors de l'étude des vibrations de réseau (chapitre 31, section 31.3).

[65]Visiblement, l'énergie de la bande pleine avec les *gaps* est plus basse que celle de la bande non déformée.

que les propriétés magnétiques d'un corps permettent de distinguer deux grandes classes de phénomènes :

- Le paramagnétisme, rendu possible lorsque les entités constitutives du solide (atomes, ions, molécules) possèdent chacune un moment magnétique non-nul en moyenne, qu'il soit d'origine orbitale, de spin ou des deux. L'application d'un champ magnétique oriente préférentiellement ces moments et une magnétisation moyenne non-nulle apparaît (paramagnétisme de Langevin[66]). Elle disparaît en champ nul, et permet de ce fait de définir une susceptibilité magnétique, χ, toujours finie à température finie (loi de Curie, $\chi \propto T^{-1}$).
- Le ferromagnétisme, qui qualifie l'aptitude d'un corps (Fe, Ni, Co,...) à conserver une aimantation macroscopique (dite *spontanée*) une fois le champ magnétique *annulé* (*aimants permanents*), à condition d'être à une température inférieure à une température remarquable (critique), appelée *température de Curie* (environ 1 000 K pour le fer). Cette propriété implique l'existence de moments microscopiques qui coopèrent, tous alignés dans le même sens sur une distance très grande à l'échelle atomique (domaines de Weiss).

Les deux phases (para- et ferromagnétique) peuvent être mieux comparées en s'appuyant sur la notion d'ordre. Un ferromagnétique est capable de maintenir un ordre à longue distance, et une aimantation spontanée une fois le champ remis à zéro. Au contraire, dans une phase paramagnétique, la magnétisation est proportionnelle au champ à petit champ, et disparaît donc avec lui, quelle que soit la température ; les fonctions de corrélation décroissent exponentiellement (ordre à courte portée), résultat d'une compétition entre la tendance à l'alignement sur le champ appliqué et l'inclination au désordre provoquée par les fluctuations thermiques. En champ nul, à l'instar d'un gaz, il n'y a pas d'ordre du tout[67]. On définit d'une façon générale un *paramètre d'ordre*, qui, pour un ferromagnétique, est nul dans la phase haute température, fini dans la phase basse température.

Bien d'autres structures magnétiques sont concevables, et ont été observées. Un cas d'une extrême richesse, et d'une grande difficulté, est l'antiferromagnétisme, qui donne lieu à des phénomènes remarquables[68] de *frustration*. C'est l'une des raisons pour

[66] ou sa version quantique, paramagnétisme de Brillouin, voir Tome I, problème 11.3.1.

[67] La question de savoir pourquoi on passe – pour certains matériaux – d'un état désordonné (paramagnétique) à un état ordonné (ferromagnétique) quand on abaisse suffisamment la température est l'enjeu de la théorie des transitions de phase. Dans ce contexte, la notion de phase apparaît comme le résultat de compétitions électroniques très subtiles, ne se traduisant pas par une modification structurale au sens premier du terme. Notons enfin que les manifestations spectaculaires du ferromagnétisme ne sont pas l'apanage de la phase solide : il existe des ferrofluides, suspensions colloïdales de microparticules magnétiques aux propriétés quelque peu magiques.

[68] D'une façon générale, l'ordre antiferromagnétique est caractérisé par une alternance régulière des moments magnétiques, dirigés tantôt vers le haut, tantôt vers le bas. Pour qu'une telle structure soit possible, il faut que l'énergie d'interaction entre deux spins premiers voisins soit *négative*, mais ce n'est pas suffisant puisque des aspects *géométriques* sont déterminants. En effet, on peut de toute évidence réaliser un tel ordre sur un réseau unidimensionnel, ou sur un réseau cubique simple, ou encore sur toute structure constituée par deux sous-réseaux imbriqués et décalés.

En revanche, on voit tout de suite qu'il est impossible de réaliser une telle disposition sur un réseau

lesquelles le cas antiferromagnétique est, en règle générale, un problème très ardu – dont la difficulté peut s'appréhender au fait que trouver le seul état fondamental du modèle de Heisenberg (défini plus loin, p. 1417) est un problème non résolu, sauf dans le cas d'une chaîne ($D = 1$) de spins $S = 1/2$ avec des couplages limités aux premiers voisins, dont la solution a été trouvée par Bethe. On se borne dans la suite au cas ferromagnétique, dont on va mettre en évidence les excitations de basse énergie appelées *magnons*.

Enfin, il n'est pas inutile de rappeler que le magnétisme – quel qu'il soit – n'est explicable que dans un cadre quantique, en raison du théorème de Miss van Leeuwen. On a vu, dans le cas le plus simple conceptuellement (l'atome d'hélium avec ses *deux* électrons), que l'*ampérien* d'Ampère est le résultat de l'action conjointe du Principe de Pauli et de la répulsion électrostatique entre les électrons, d'où l'ordre de grandeur du *splitting* singulet - triplet, incompréhensible au vu de la petitesse des interactions magnétiques, au sens classique du terme.

30.6.1 Définition des magnons (ferromagnétiques)

Le théorème de Bloch n'est pas seulement utile pour étudier les propriétés électroniques d'un cristal. S'appuyant sur la symétrie de translation discrète, il ressort en fait à propos de toute excitation élémentaire, qu'il s'agisse de vibrations de réseau (*phonons*, voir chapitre 31), excitons (voir note 32, chapitre 32) ou de n'importe quel autre type de degré de liberté, magnétique par exemple.

La notion de magnon est facile à appréhender, dans le droit fil des idées énoncées plus haut. Supposons donné un réseau de Bravais de N sites, décoré avec des atomes ayant chacun un moment magnétique, appelé simplement *spin* dans toute la suite, ayant la valeur $S = 1/2$, cette valeur précise n'étant pas anodine : on peut renverser un tel spin par l'opérateur S_-, mais pas plus : $S_-^2|\uparrow\rangle = 0$ – ceci simplifie considérablement le problème. On se place d'emblée dans le cas où l'interaction entre spins est essentiellement *attractive*, ce qui injecte dans le problème une tendance naturelle au ferromagnétisme : à température nulle, tous les spins veulent être alignés, c'est l'état ferromagnétique pur. Cet état est le fondamental (le *vide*) du cristal, tant que l'on ne considère que ses degrés de liberté de spin ; il est caractérisé par le spin maximum $S_{\text{max}} = N/2$, $M_S = N/2$ (tous les spins pointent "en haut", conventionnellement). Enfin, la longueur a désignera la dimension linéaire typique d'une maille primitive.

Savoir ce qu'est l'état fondamental d'un système est une information importante, mais elle ne permet pas, à elle seule, d'en étudier les propriétés thermodynamiques, qui exigent de connaître les états excités, au moins les premiers d'entre eux tant que la température n'est pas trop élevée. Une question majeure en la matière est l'existence ou non d'un *gap* d'énergie entre l'état fondamental et les premières excitations : de

triangulaire ; le troisième spin a envie d'être en haut quand il regarde d'un côté, et en bas quand il regarde de l'autre, d'où la notion de *frustration* : finalement, que faire ? C'est un cas où, toute autre interaction mise à part, l'état fondamental est *infiniment* dégénéré (il existe alors une entropie résiduelle non nulle à température nulle).

toute évidence, la chaleur spécifique c_V aura selon les cas un comportement radicalement différent. S'il y a un *gap*, c_V s'effondre exponentiellement vite à basse température ($c_V \propto \mathrm{e}^{-A/T}$), c'est le cas pour un gaz parfait de spins en champ magnétique. Au contraire, s'il existe des excitations d'énergie arbitrairement petite, $c_V \propto T$ (par exemple : gaz d'électrons dégénéré). D'une façon générale, l'existence d'un *gap* d'énergie E_g donne une grande robustesse à l'état fondamental, dont la pertinence est préservée à température finie, petite devant $E_\mathrm{g}/k_\mathrm{B}$. En l'absence de *gap*, au contraire, l'état fondamental est un état idéal dont le contenu physique est *a priori* très limité.

Il importe donc d'étudier (au moins) les premiers états excités du cristal ferromagnétique. On voit de suite qu'une façon d'exciter le système au-dessus du fondamental est de basculer le spin[69] de l'un des atomes, ce qui a un coût énergétique puisqu'il existe un état fondamental ordonné, seulement possible si les spins interagissent entre eux (et ici de façon attractive). Ceci réalisé, on n'a que l'embarras du choix : de quel atome va-t-on retourner le spin ? En fait, tous les états où un seul spin a basculé ont exactement la même énergie, de sorte que, à ce stade, on dispose de $N \gg 1$ états (d'ordre zéro) (une infinité en quelque sorte), tous dégénérés, engendrant un sous-espace $\mathcal{E}_1$.

Désignons par $\vec{R} \in \mathcal{B}$ le rayon-vecteur du site dont le spin $\vec{S}_{\vec{R}}$ a été basculé (il pointe en bas, avec la convention adoptée), et par $|\vec{R}\rangle$ l'état de spin correspondant ; cet état n'est sûrement pas propre du Hamiltonien H – pas encore écrit d'ailleurs – puisque, les spins interagissant, retourner l'un d'entre eux ne saurait laisser les autres indifférents. Quel que soit le Hamiltonien modèle choisi par la suite pour les calculs, il est certain qu'il possédera la symétrie de translation du réseau : il commute donc avec tous les opérateurs de translation $\mathcal{T}(\vec{R})$. Tout comme dans le schéma des liaisons fortes, on dispose donc de N états *localisés* au sens où chacun d'entre eux, par définition, porte un *défaut* ponctuel (le spin retourné) ; ils sont tous dégénérés et tous se transforment les uns dans les autres par les opérateurs $\mathcal{T}(\vec{R})$. En répétant exactement le même argument que celui faisant passer de (30.149) à (30.154), on obtient l'état noté simplement $|\vec{k}\rangle$:

$$\boxed{|\vec{k}\rangle \stackrel{\text{déf}}{=} \frac{1}{\sqrt{N}} \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\, |\vec{R}\rangle} \tag{30.201}$$

qui est propre de tous les $\mathcal{T}(\vec{R})$, et satisfait le théorème de Bloch, au même titre que l'état $|\psi_{\vec{k}}\rangle$ défini en (30.154) (ici, la constante de normalisation vaut simplement $N^{-1/2}$ puisque tous les états localisés $|\vec{R}\rangle$ sont visiblement orthogonaux). Satisfaisant le théorème de Bloch, et en l'absence d'autres degrés de liberté (autres nombres quantiques), cet état $|\vec{k}\rangle$ est aussi propre du Hamiltonien (rappelons que l'aspect *abélien* du groupe des translations joue un rôle important pour cette propriété de $|\vec{k}\rangle$). L'excitation représentée par l'état $|\vec{k}\rangle$ ci-dessus est appelée *magnon*, et est caractérisée par l'impulsion $\hbar\vec{k}$; noter qu'elle est équirépartie sur le réseau : la probabilité de trouver que le spin basculé est en un certain site $\vec{R}_0$ vaut $1/N$: elle est indépendante de $\vec{R}_0$ et est donc la même partout.

S'agissant de pousser les calculs plus avant, le choix d'un Hamiltonien-modèle H s'impose. Indépendamment de la symétrie de translation liée au réseau, et en l'absence

[69]Dans le jargon courant, quand un spin a été retourné, on dit qu'il a *flippé*.

de toute direction privilégiée, H doit également être invariant dans toute rotation de l'ensemble des spins. La quantité la plus simple ayant cette propriété est un produit scalaire, par exemple $\vec{S}_{\vec{R}}.\vec{S}_{\vec{R}'}$. Pour avoir l'invariance de translation, il suffit de sommer sur le réseau ; quant à l'interaction entre deux spins $\vec{S}_{\vec{R}}$ et $\vec{S}_{\vec{R}'}$, elle ne doit dépendre que de $\vec{R}-\vec{R}'$, justement pour avoir un systène invariant de translation. En définitive, le Hamiltonien-modèle le plus simple a l'expression[70] :

$$\boxed{H \stackrel{\text{déf}}{=} -\frac{1}{2\hbar^2}\sum_{(\vec{R},\,\vec{R}')\in\mathcal{B}^2} J(\vec{R}-\vec{R}')\,\vec{S}_{\vec{R}}.\vec{S}_{\vec{R}'}} \tag{30.202}$$

et s'appelle le *Hamiltonien de Heisenberg.* L'interaction entre deux spins est décrite par la fonction $J(\vec{R}-\vec{R}')$, pour l'instant quelconque et homogène à une énergie, mais comme $J(\vec{R})$ est ici censée exprimer la tendance au ferromagnétisme, cette fonction doit être positive pour les valeurs de $\vec{R}$ les plus pertinentes. Deux opérateurs de spin sur deux sites différents commutent entre eux, ce qui entraîne que $J(\vec{R})$ est une fonction paire (on permute les opérateurs) et réelle (H doit être hermitique). Enfin, l'état où tous les spins sont alignés est visiblement propre de H.

L'énergie de l'état fondamental $|\text{vac}\rangle$ est :

$$E_{\text{vac}} = \langle\text{vac}|H|\text{vac}\rangle = -\frac{1}{2\hbar^2}\sum_{\vec{R},\vec{R}'} J(\vec{R}-\vec{R}')\big(\frac{\hbar}{2}\big)^2 = -\frac{N}{8}\sum_{\vec{R}} J(\vec{R}) \ . \tag{30.203}$$

Le sous-espace $\mathcal{E}_1$ (un seul spin retourné) est visiblement stable par H ; comme de surcroît $\langle\vec{k}|H|\vec{k}'\rangle = 0$ si $\vec{k}\neq\vec{k}'$, il suffit de calculer l'élément diagonal pour avoir la valeur propre, c'est-à-dire la loi de dispersion $\varepsilon(\vec{k}\,)$ d'un magnon. En utilisant, comme pour les liaisons fortes, l'invariance par translation du Hamiltonien (voir le passage de (30.170) à (30.171)), on a :

$$\varepsilon(\vec{k}\,) = \langle\vec{k}|H|\vec{k}\rangle = \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\langle\vec{0}|H|\vec{R}\rangle \ . \tag{30.204}$$

Le calcul de l'élément de matrice est un peu laborieux, mais ne pose pas de réelle difficulté. On utilise d'une part :

$$\vec{S}_{\vec{R}}.\vec{S}_{\vec{R}'} = S_{\vec{R}z}S_{\vec{R}'z} + \frac{1}{2}(S_{\vec{R}+}S_{\vec{R}'-} + S_{\vec{R}-}S_{\vec{R}'+}) \equiv S_{\vec{R}z}S_{\vec{R}'z} + \vec{S}_{\vec{R}\perp}.\vec{S}_{\vec{R}'\perp} \ , \tag{30.205}$$

d'autre part les égalités :

$$S_{\vec{R}'z}S_{\vec{R}''z}|\vec{R}\rangle = \frac{\hbar^2}{4}(1-2\delta_{\vec{R}\vec{R}'})(1-2\delta_{\vec{R}\vec{R}''})|\vec{R}\rangle \ , \tag{30.206}$$

[70] Comme annoncé dans le chapitre 27 p. 1224 (et problème 27.8.1), le Hamiltonien de Heisenberg est une généralisation pour N spins du terme de couplage effectif spin-spin qui apparaît naturellement pour un système de deux spins.

$$S_{\vec{R}'+}S_{\vec{R}''-}|\vec{R}\rangle = \hbar^2(1-\delta_{\vec{R}\vec{R}''})(\delta_{\vec{R}'\vec{R}''}|\vec{R}\rangle + \delta_{\vec{R}\vec{R}'}|\vec{R}''\rangle) \;; \tag{30.207}$$

(30.207) provient de l'application répétée de $S_{\vec{R}'-}|\vec{R}\rangle = (1-\delta_{\vec{R}\vec{R}'})|\vec{R}\vec{R}'\rangle$; on trouve finalement :

$$\boxed{\varepsilon(\vec{k}\,) = E_{\text{vac}} + \sum_{\vec{R}\in\mathcal{B}} J(\vec{R})\,\sin^2\frac{\vec{k}.\vec{R}}{2}} \tag{30.208}$$

Ce résultat donne la relation de dispersion d'un magnon, et apporte la réponse à l'une questions importantes évoquées plus haut : il n'y a pas de *gap* entre le fondamental et les premières excitations[71], dont celles de plus basse énergie correspondent à $\|\vec{k}\|$ très petit. Dans cette limite, on a :

$$\|\vec{k}\| \ll \frac{\pi}{a} \;:\; \varepsilon(\vec{k}\,) - E_{\text{vac}} \simeq \frac{1}{4}\sum_{\vec{R}\in\mathcal{B}} J(\vec{R})\,(\vec{k}.\vec{R})^2 \;. \tag{30.209}$$

Ces excitations démarrent donc typiquement comme k^2 pour les grandes longueurs d'onde (ou dans la limite d'un réseau quasi-continu spatialement) : non seulement il n'y a pas de *gap*, mais en plus l'énergie varie quadratiquement en k. En y allant très rapidement par un simple argument de *scaling*, on voit que l'énergie moyenne à la température T variera comme[72] $T^{1+D/2}$ pour un réseau dans $\mathbb{R}^D$, ce qui donne $c_V \propto T^{D/2}$ (comparer avec les phonons acoustiques, éq. (31.114)).

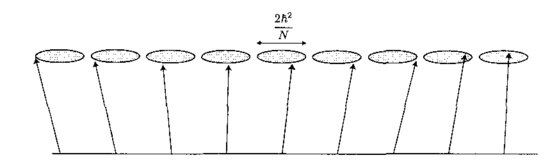


Figure 30.10: Illustration schématique de la précession spatiale de la magnétisation transverse d'un magnon.

Une autre quantité intéressante est la fonction de corrélation transverse dans l'état de magnon :

$$C_{\vec{k}}(\vec{R},\,\vec{R}') \stackrel{\text{déf}}{=} \hbar^{-2}\langle\vec{k}|\vec{S}_{\vec{R}\perp}.\vec{S}_{\vec{R}'\perp}|\vec{k}\rangle \;; \tag{30.210}$$

on trouve sans difficulté :

$$\boxed{C_{\vec{k}}(\vec{R},\,\vec{R}') = \frac{1}{2}\delta_{\vec{R}\vec{R}'} + \frac{1}{N}(1-\delta_{\vec{R}\vec{R}'})\,\cos\vec{k}.(\vec{R}-\vec{R}')} \tag{30.211}$$

[71] Noter que $\varepsilon(\vec{k}\,)$ donné en (30.208) possède la périodicité caractéristique d'une excitation sur réseau : $\varepsilon(\vec{k}+\vec{K}) = \varepsilon(\vec{k}\,)$, $\forall\,\vec{K}\in\mathcal{B}$.

[72] Pour une loi de dispersion variant comme k^α à petit k, et sans autre type d'excitations de basse énergie, l'énergie est $\propto T^{1+(D/\alpha)}$ à basse températeure (voir exercice 30.7.7) : avec $\alpha = 1$ (cas des phonons acoustiques), on a $c_V \propto T^3$ pour $D = 3$ (voir chapitre 31, section 31.6). En l'absence de *gap*, un comportement en puissance de la loi de dispersion produit une variation de même nature pour c_V, *a priori* d'autant plus rapide que la courbe de dispersion démarre lentement à partir de $k = 0$: plus la relation de dispersion est "écrasée", plus les excitations sont importantes, plus vite augmente l'énergie moyenne (et la chaleur spécifique).

Ce résultat signifie qu'il existe une polarisation *transverse* (toute petite) ; à partir d'un site de référence ($\vec{R} = \vec{0}$), elle tourne spatialement (précession dans l'espace), le déphasage étant donné par $\cos \vec{k}.\vec{R}$ (voir fig. 30.10).

30.6.2 États liés de deux magnons

On vient de voir que les magnons apparaissent tout naturellement quand on bascule un spin, et constituent donc les *premières* excitations du modèle de Heisenberg ferromagnétique. La question arrive maintenant : que se passe-t-il si on bascule non pas *un* mais *deux* spins ? La réponse est loin d'être triviale, et on y répond dans la suite pour un système beaucoup plus simple : un réseau à une dimension, de pas a, et avec des interactions $J(\vec{R})$ limitées aux premiers voisins. On verra qu'il existe des *paires liées*, bel exemple d'association de deux excitations libres qui, grâce à une interaction effective *attractive*, peuvent former des *molécules* d'un nouveau genre.

Avec les hypothèses simplificatrices juste énoncées, et en présence en plus d'un champ magnétique homogène $\mathcal{B}$ orienté le long de Oz, le Hamiltonien ferromagnétique de Heisenberg a pour expression :

$$H = -\hbar^{-2} J \sum_{n=1}^{N} \vec{S}_n.\vec{S}_{n+1} - \hbar^{-1}\gamma\mathcal{B}\sum_{n=1}^{N} S_{zn} \ . \tag{30.212}$$

l'indice n numérote les sites ($N + 1 \equiv 1$), $\gamma = g\mu_{\mathrm{B}}$ est le facteur gyromagnétique, supposé positif ; comme $S = \frac{1}{2}$, le couplage Zeeman d'un spin avec le champ est $\hbar^{-1}\gamma\mathcal{B}\frac{\hbar}{2} = \frac{1}{2}\gamma\mathcal{B}$: basculer un spin augmente de $\gamma\mathcal{B}$ l'énergie Zeeman. J mesure l'énergie d'interaction de deux spins voisins, positive puisqu'il est toujours question de ferromagnétisme. Dans l'état fondamental, conventionnellement, tous les spins pointent "en haut" ; chacun d'entre eux interagit avec ses deux voisins, l'énergie étant alors à chaque fois $\hbar^{-2}J(\frac{\hbar}{2})^2 = \frac{J}{4}$, soit $\frac{J}{2}$ au total – d'où le coût $+J$ pour retourner *un* spin. L'énergie fondamentale est donc $-\frac{1}{2}N\frac{J}{2} = -N\frac{J}{4}$, en accord avec (30.203), la somme $\sum_{\vec{R}} J(\vec{R})$ valant ici $J(-a) + J(+a) = 2J$. Noter que le basculement d'un spin fait varier de $\hbar$ la composante S_z du spin total : considérant le fondamental comme l'état vide de magnons, le magnon apparaît comme une particule de spin 1.

Pour ce réseau unidimensionnel, l'état de magnon $|k\rangle$ s'écrit :

$$|k\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \mathrm{e}^{\mathrm{i}nka}|n\rangle \tag{30.213}$$

(comparer avec (30.201)), où $|n\rangle$ est l'état localisé correspondant au spin basculé sur le site n. Pour le magnon de centre de zone ($k = 0$), tous les coefficients valent 1 : l'état $|k = 0\rangle$ est invariant dans toute translation, $\mathcal{T}(a)|k = 0\rangle = |k = 0\rangle$, en conformité avec le fait qu'il a une énergie égale à celle du fondamental. À l'opposé, pour le magnon de bord de zone, $k = \pm\frac{\pi}{a}$, les coefficients sont alternés et l'état change de signe par la translation

élémentaire : $\mathcal{T}(a)|k = \pm\frac{\pi}{a}\rangle = -|k = \pm\frac{\pi}{a}\rangle$. Comme dans tous les cas les poids de tous les sites sont les mêmes, on s'attend à ce que l'alternance des poids donne à l'énergie du magnon sa valeur maximum ; c'est bien le cas (voir ci-dessous) – d'ailleurs, comme l'état $|k = \pm\frac{\pi}{a}\rangle$ ne change pas par les translations de $2qa$ ($q \in \mathbb{Z}$), on imagine aisément une structure moyenne du genre antiferromagnétique, qui coûte beaucoup d'énergie. Noter que $|n\rangle \equiv S_{n-}|\text{vac}\rangle$, de sorte que l'on a aussi :

$$|k\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \mathrm{e}^{inka} S_{n-}|\text{vac}\rangle \ . \tag{30.214}$$

Quant aux énergies, elles sont (voir (30.203) et (30.208)) :

$$E_{\text{vac}} = -\frac{NJ}{4} - N\frac{\gamma\mathcal{B}}{2} \ , \qquad \varepsilon(k) = E_{\text{vac}} + 2J\sin^2\frac{ka}{2} + \gamma\mathcal{B} \ , \tag{30.215}$$

d'où l'énergie d'un magnon en champ nul :

$$\boxed{\varepsilon_{\text{magnon}}(k) \overset{\text{déf}}{=} \varepsilon(k) - E_{\text{vac}} = 2J\sin^2\frac{ka}{2} = J(1-\cos ka)} \tag{30.216}$$

l'énergie maximum, égale à $2J$, est bien en bord de zone ($ka = \pi$), et est la même que si on avait retourné deux spins distants ; la loi de dispersion à petit k est $\varepsilon_{\text{magnon}}(k) \simeq \frac{1}{2}k^2a^2J$: la masse effective d'un magnon de petit k est $m_{\text{eff}} = \frac{\hbar^2}{Ja^2}$. La figure 30.11 illustre la situation $D = 1$.

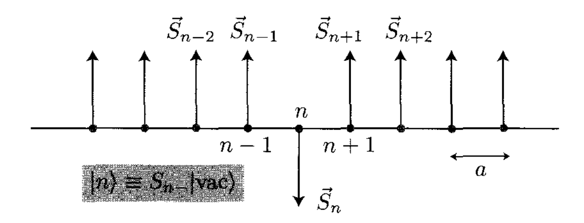


Figure 30.11: Réseau unidimensionnel avec son spin retourné au site n, définissant l'état localisé d'ordre zéro $|n\rangle$, d'énergie $+J$ au-dessus du fondamental.

Avant d'attaquer le problème des doubles excitations, un résultat mérite d'être noté, qui s'établit facilement dans la version unidimensionnelle. Supposons qu'il est possible (par collision inélastique de neutrons par exemple) de préparer un état initial où seul le spin $\vec{S}_n$ est basculé vers le bas. Cet état n'étant pas stationnaire, il évolue de façon non triviale, mais il est facile de trouver l'état à l'instant t ultérieur et d'obtenir, par exemple, la valeur moyenne du spin $\vec{S}_{n'}$ à cet instant. On trouve notamment :

$$\hbar^{-1}\langle S_{zn'}\rangle(t) = \frac{1}{2} - J_{|n-n'|}(Jt/\hbar) \ , \tag{30.217}$$

où J_n est une fonction de Bessel ordinaire ; on sait que c'est une fonction oscillante (ici au cours du temps), avec une amplitude décroissant comme $t^{-1/2}$: l'excitation initialement

localisée se dilue lentement le long de la chaîne, en oscillant, et ne présente aucune forme de périodicité (ou même quasi-périodicité). Un tel comportement est assez universel ; on le retrouvera à l'identique à propos des vibrations de réseau (voir (31.53)), où il décrira alors la relaxation du champ des déplacements d'une chaîne d'atomes après qu'un atome a été initialement écarté de sa position d'équilibre.

Il s'agit maintenant d'examiner les excitations se situant dans le sous-espace $\mathcal{E}_2$ engendré par les vecteurs où deux spins sont renversés. Pour résoudre ce problème en 1931, Bethe [247] a inventé une méthode connue depuis sous le nom de *Bethe Ansatz*, qui joue un grand rôle en permettant de résoudre des problèmes très difficiles de particules en interaction. En règle générale, l'*Ansatz* de Bethe, qui ne fonctionne que pour certains modèles bien particuliers[73], est un exercice de virtuosité algébrique[74], et produit des solutions dont il est souvent difficile d'extraire de l'information. Ici, le niveau de technicité requis est très raisonnable, comme on pourra en juger, et constitue une introduction très accessible à une technique souvent difficile à conduire jusqu'au bout, mais importante.

Étant donné que l'on ne peut pas exciter deux fois le même spin, l'espace de travail $\mathcal{E}_2$ est de dimension $\frac{1}{2}N(N-1)$. $\mathcal{E}_2$ n'est pas couplé au sous-espace $\mathcal{E}_1$ à un magnon (qui est propre de H), ni à l'état fondamental (également propre). Enfin, comme H ne peut basculer que deux spins à la fois, $\mathcal{E}_2$ n'est pas non plus couplé aux sous-espaces $\mathcal{E}_{p>2}$ correspondants au retournement de p spins. Au total, tout comme $\mathcal{E}_1$, l'espace $\mathcal{E}_2$ est lui aussi *stable* par H.

D'un autre côté, on peut formellement définir des états $|k\rangle \otimes |k'\rangle \equiv |kk'\rangle$, dont l'énergie est :

$$\varepsilon_{\text{deux magnons libres}} \stackrel{\text{déf}}{=} \varepsilon_{\text{magnon}}(k) + \varepsilon_{\text{magnon}}(k') = J(2 - \cos ka - \cos k'a) \quad . \tag{30.218}$$

Il y a $\frac{1}{2}N(N-1)$ états avec $\vec{k} \neq \vec{k}'$, et N autres avec $\vec{k} = \vec{k}'$; au total, il existe ainsi $\frac{1}{2}N(N+1)$ états à deux magnons libres. En nombre supérieur à la dimension du sous-espace $\mathcal{E}_2$, ils ne peuvent être deux à deux orthogonaux et, ne serait-ce que pour cette raison, ne peuvent être propres du Hamiltonien. Ainsi, les états à deux magnons libres $|k, k'\rangle$ ne sont pas des excitations stationnaires, et évoluent au cours du temps – sauf bien sûr si k ou k' est nul. On peut se faire une image dynamique de cette situation en disant que les deux magnons se propagent, entrent en collision et repartent chacun de son côté avant une nouvelle collision, et ainsi de suite. La notion de *collision* est d'autant plus justifiée que la somme $k + k'$ est une constante du mouvement, comme on va le voir ; les collisions entre deux magnons se font donc en respectant la loi de conservation :

$$\boxed{k_{\text{i}} + k'_{\text{i}} = k_{\text{f}} + k'_{\text{f}}} \tag{30.219}$$

[73]Par exemple, le modèle de Hubbard $D = 1$ sur réseau rigide, résolu par Lieb et Wu en 1968 [241]. Cette méthode fonctionne aussi pour des bosons à une dimension en interaction de contact. Ici, l'interaction de deux spins retournés n'est différente de zéro que s'ils sont sur deux sites premiers voisins, ce qui correspond à leur plus petite distance possible, puisque on ne peut pas basculer davantage un spin déjà retourné. En ce sens, on peut aussi parler d'interaction *de contact*, et c'est bien pourquoi l'*Ansatz* de Bethe est efficace dans le type de modèle évoqué, et pour les magnons en interaction, à condition que $S = 1/2$. Il s'agit en fait d'un exemple d'application de la version *discrète* (sur réseau) de l'*Ansatz*, par opposition à des particules en interaction suivant $\delta(x_i - x_j)$.

[74]L'ouvrage de référence en la matière est celui de Michel Gaudin [177].

Le fait que $|kk'\rangle$ n'est pas stationnaire peut se comprendre en réalisant les tendances antagonistes du problème : quand ils sont côte-à-côte, les deux spins renversés sont en interaction attractive l'un avec l'autre, mais dès qu'ils s'éloignent, les autres spins ont tendance à vouloir les redresser. Le résultat net de cette compétition est élucidé dans ce qui suit.

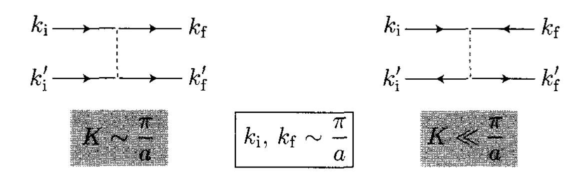


Figure 30.12: Conservation de l'impulsion lors de la collision de deux magnons libres. La figure de droite montre un *rebond* de chaque magnon sur l'autre.

La conservation de la somme $\vec{k}+\vec{k}'$ s'établit comme suit. Par construction, chaque état $|k\rangle$ est propre des opérateurs de translation $\mathcal{T}(qa)$ ($q \in \mathbb{Z}$) ; on a en particulier :

$$\mathcal{T}(a)|k\rangle = \mathrm{e}^{-ika}|k\rangle \ , \tag{30.220}$$

une égalité qui se vérifie immédiatement en utilisant la définition habituelle de l'opérateur de translation dans l'espace des états, $\mathcal{T}(a)|n\rangle \overset{\text{déf}}{=} |n+1\rangle$, et qui est la version unidimensionnelle de (30.22). L'état à deux magnons libres $|kk'\rangle$ satisfait donc :

$$\mathcal{T}(a)|kk'\rangle = \mathrm{e}^{-\mathrm{i}(k+k')a}|kk'\rangle \ , \tag{30.221}$$

montrant que cet état est propre de $\mathcal{T}(a)$ et associé à la valeur propre $\mathrm{e}^{-\mathrm{i}(k+k')a}$. Comme H commute avec $\mathcal{T}(a)$, les éléments de matrice de H entre deux états propres de $\mathcal{T}(a)$ associés à deux valeurs propres *distinctes* de $\mathcal{T}(a)$ sont nuls : H ne peut donc transformer k et k' qu'en respectant la valeur de leur somme – c'est ce qu'exprime (30.219).

Il s'agit donc de trouver les états propres $|\psi\rangle$ de H dans le sous-espace $\mathcal{E}_2$ stable par H. On les écrit naturellement sous la forme :

$$|\psi\rangle = \sum_{n,n'} c_{nn'}|nn'\rangle \equiv \sum_{n,n'} c_{nn'}S_{n-}S_{n'-}|\text{vac}\rangle \ , \tag{30.222}$$

$|nn'\rangle$ désignant le vecteur de base de $\mathcal{E}_2$ où les spins inversés sont sur les deux sites d'indices n et n'. La question maintenant est de trouver les coefficients $c_{nn'} = c_{n'n}$, cette symétrie résultant de la commutation $[\vec{S}_n,\, \vec{S}_{n'}] = 0$, $n \neq n'$ – noter aussi que $c_{nn} = 0$ en vertu de[75] $S_{n-}S_{n-}|\text{vac}\rangle = 0$. Il est quelque peu laborieux, mais sans réelle difficulté, de trouver la relation entre ces coefficients : on calcule l'effet de H sur un vecteur de base $|nn'\rangle$ et on identifie les composantes des vecteurs d'un membre à l'autre de $H|\psi\rangle = E|\psi\rangle$ (les calculs détaillés sont proposés dans le problème 30.7.10). On trouve ainsi :

$$\boxed{\lambda c_{nn'} + \frac{1}{2}(c_{nn'-1} + c_{nn'+1} + c_{n-1n'} + c_{n+1n'}) = -(\delta_{n'n+1} + \delta_{n'n-1})} \tag{30.223}$$

[75] Ici joue à plein l'hypothèse de spins 1/2.

où λ donne l'énergie E par l'égalité[76] :

$$E = E_{\text{vac}} + 2\gamma\mathcal{B} + 2J + \lambda J \ . \tag{30.224}$$

Le terme $+2\gamma\mathcal{B}$ se comprend bien : c'est juste le surcoût énergétique Zeeman des deux spins retournés, où qu'ils se trouvent ($\mathcal{B}$ est homogène). Le terme additif $2J$ vient du basculement des deux spins, chacun en interaction avec son premier voisin ; le dernier terme contient finalement toute l'inconnue du problème en cours.

Si le second membre de la récurrence (30.223) était toujours nul, la solution serait $c_{nn'} = \mathrm{e}^{\mathrm{i}kna}\mathrm{e}^{\mathrm{i}k'n'a} + \mathrm{e}^{\mathrm{i}kn'a}\mathrm{e}^{\mathrm{i}k'na}$, avec $\lambda = -(\cos ka + \cos k'a) \stackrel{\text{déf}}{=} \lambda_K^{(0)}(q)$, donnant pour l'énergie d'excitation $2J - J(\cos ka + \cos k'a)$, qui est juste la somme des énergies des deux magnons $\varepsilon(k) + \varepsilon(k')$ (voir (30.216)). Un état du genre $|kk'\rangle$ serait alors propre de H – mais on sait bien que ce n'est pas le cas – et chaque coefficient serait caractérisé par deux valeurs k et k', chacune fixée une fois pour toutes.

Ce second membre représente de fait l'interaction entre deux spins retournés quand ils sont côte à côte, qui produit l'image dynamique de magnons libres interagissant lors de collisions répétées comme indiqué plus haut, entraînant que k et k' ne sont pas séparément des constantes du mouvement. Comme la somme $k+k'$ est conservée lors de ces collisions, le coefficient $c_{nn'}$ peut être cherché sous la forme d'une somme de Fourier sur k et k' avec la contrainte $k+k' = \mathrm{C^{ste}}$; par ailleurs, il est physiquement utile d'introduire l'impulsion relative des deux magnons, d'où les définitions[77] naturelles :

$$K \stackrel{\text{déf}}{=} k + k' \ , \quad q \stackrel{\text{déf}}{=} \frac{1}{2}(k - k') \quad \Longleftrightarrow \quad k = \frac{K}{2} + q \ , \quad k' = \frac{K}{2} - q \ . \tag{30.226}$$

L'impulsion totale $\hbar K$ varie *a priori* entre $\pm\hbar\frac{2\pi}{a}$ quand k et k' sont pris dans la première zone de Brillouin , mais comme ces nombres d'onde sont eux-mêmes définis à $\frac{2\hbar\pi}{a}$ près, il en va de même pour K, que l'on peut donc également prendre dans BZ1 sans perdre des solutions, soit $|K| \leq \pi/a$; $\hbar q$ est l'impulsion du mouvement fictif réduit. Avec ces définitions, $\cos ka + \cos k'a = 2\cos qa \cos\frac{Ka}{2}$: l'énergie de deux magnons sans interaction est $2J(1 - \cos qa \cos\frac{Ka}{2})$ (voir (30.216)) ; l'impulsion $\hbar K$ étant fixée, elle prend sa plus petite valeur[78] $2J(1 - \cos\frac{Ka}{2})$ pour $q = 0$, c'est-à-dire quand ils ont la même impulsion.

On cherche donc maintenant les coeffcients $c_{nn'}$ sous la forme :

$$c_{nn'} = \sum_q \Phi_K(q)\mathrm{e}^{\mathrm{i}(\frac{K}{2}+q)na}\,\mathrm{e}^{\mathrm{i}(\frac{K}{2}-q)n'a} \ . \tag{30.227}$$

---

[76]Noter que le champ magnétique n'intervient pas dans cette récurrence, puisqu'il figure dans H en tant que terme à *un seul* corps.

[77]Attention : K désigne ici l'impulsion totale, pas un vecteur du réseau réciproque !

Noter que le changement de variables (30.225) correspond à la réduction d'un problème à deux corps en interaction centrale. Par dualité, la transformation dans l'espace réel est :

$$R \stackrel{\text{déf}}{=} \frac{1}{2}(r + r') \ , \quad \rho \stackrel{\text{déf}}{=} r - r' \quad \Longleftrightarrow \quad r = R + \frac{\rho}{2} \ , \quad r' = R - \frac{\rho}{2} \ , \tag{30.225}$$

où r et r' sont les abscisses des deux spins renversés. Les deux particules étant identiques, la masse réduite du magnon *fictif* est la moitié de la masse de chacune d'entre elles.

[78]Comme $|K| \leq \frac{\pi}{a}$, $\cos\frac{Ka}{2}$ est toujours positif.

La fonction $\Phi_K(q)$ est une sorte de fonction d'onde en représentation-p, paire en q pour assurer la symétrie $c_{nn'} = c_{n'n}$, et satisfaisant en outre $\sum_q \Phi_K(q) = 0$ puisque $c_{nn} = 0$. Le report de cette forme de $c_{nn'}$ dans la relation de récurrence (30.223) permet d'obtenir l'équation satisfaite par $\Phi_K(q)$; après calcul, on trouve :

$$(\lambda + 2\cos qa \cos \frac{Ka}{2})\Phi_K(q) = \frac{2}{N}\sum_{q'} \big(\cos \frac{Ka}{2} - \cos qa\big) \cos q'a \, \Phi_K(q') \ . \qquad (30.228)$$

Remarquer que la parenthèse à gauche n'est autre que $\lambda - \lambda_K^{(0)}(q)$: toute l'interaction est représentée par le second membre. Celui-ci est de la forme $\big(\cos \frac{Ka}{2} - \cos qa\big) S_K$ où la quantité S_K :

$$S_K \stackrel{\text{déf}}{=} \frac{2}{N}\sum_{q'} \cos q'a \, \Phi_K(q') \qquad (30.229)$$

est pour l'instant inconnue ; ceci permet d'écrire *formellement* la solution :

$$\Phi_K(q) = \frac{\cos \frac{Ka}{2} - \cos qa}{\lambda + 2\cos qa \cos \frac{Ka}{2}} S_K \ . \qquad (30.230)$$

Reportant cette expression de $\Phi_K(q)$ dans l'égalité de définition (30.229) de S_K, on obtient :

$$\boxed{1 = \frac{2}{N}\sum_{q} \cos qa \, \frac{\cos \frac{Ka}{2} - \cos qa}{\lambda + 2\cos qa \cos \frac{Ka}{2}} \stackrel{\text{déf}}{=} F_{NK}(\lambda)} \qquad (30.231)$$

La somme sur q porte sur toutes les valeurs $p\frac{2\pi}{N}$ avec N valeurs distinctes pour l'entier p (par exemple : $p = 0, 1, 2, ..., N-1$) – on note au passage que cette équation ne change pas si Ka sort de l'intervalle $[-\pi, +\pi]$ puisque, dans la somme, la variable muette qa décrit une période 2π complète. K étant fixé, le second membre est une fonction de la seule variable λ, valeur propre réduite en unités J. Il en résulte que (30.231) constitue de fait une équation pour λ : c'est un avatar de l'équation aux valeurs propres.

Cette forme de l'équation donnant les valeurs propres a déjà été rencontré à de multiples reprises[79] – voir par exemple la discussion à propos du polyacétylène, chapitre 28, sous-section 28.2.5, fig. 28.16. Ici encore, l'analyse passe par une étude graphique préalable de cette équation, en traçant le graphe de $F_{NK}(\lambda)$, et en la coupant par la droite horizontale de cote égale à 1. Cette fonction diverge pour toutes les valeurs de λ égales à $-2\cos p\frac{2\pi}{N} \cos \frac{Ka}{2}$, et à condition que $p \neq p_0 \stackrel{\text{déf}}{=} \frac{NK}{4\pi}$. Une étude analytique un peu détaillée permet de reconstituer tous les détails qualitatifs visibles sur la figure 30.13.

De la discussion graphique, il ressort deux conclusions importantes :

[79]De fait, cette manière de procéder pour obtenir une équation fixant les valeurs propres est assez universelle ; on la retrouvera à propos de supraconductivité (chapitre 32, sous-section 32.5.2) ; un autre exemple a été rencontré dans le Tome I, problème 14.6.11, à propos de l'allongement du temps de retour lors de la densification des états d'un système quantique, tout comme dans le problème 17.8.6.

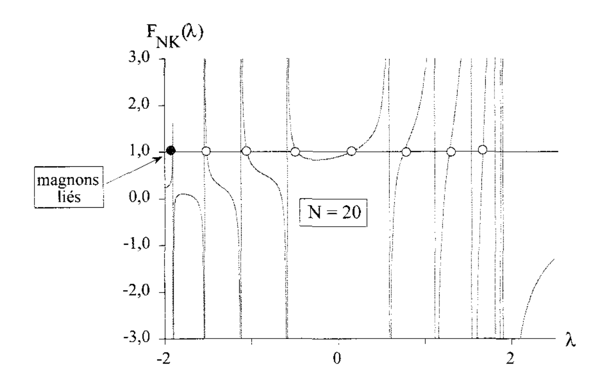


Figure 30.13: Résolution graphique de l'équation (30.231). L'état de la paire liée de deux magnons est à gauche. La figure est tracée avec $Ka = \frac{\pi}{5}$, et pour une petite valeur de N pour la lisibilité (en réalité $N \gg 1$).

1. il existe toujours une et une seule solution $\lambda_{\mathrm{b}K}$ strictement *inférieure* à[80] $-2\cos\frac{Ka}{2}$ (signalée par la flèche à gauche dans la figure 30.13), donc toujours plus basse (voir (30.218)) que la plus petite énergie de deux magnons sans interaction dont l'impulsion du centre de masse est fixée à la valeur $\hbar K$.

2. Il existe en outre un ensemble très dense de solutions, strictement comprises dans l'intervalle $\pm 2\cos\frac{Ka}{2}$, très voisines de $\lambda_K^{(0)}(q) = -2\cos\frac{Ka}{2}\cos qa$, énergie de deux magnons sans interaction. Dans la limite $N \gg 1$, les racines exactes sont quasiment indiscernables de ces valeurs explicites, et forment un quasi-continuum.

En résumé, le spectre des excitations $\varepsilon \stackrel{\text{déf}}{=} E - E_{\text{vac}}$ de H dans le sous-espace $\mathcal{E}_2$ est l'ensemble suivant, paramétré par K :

$$\varepsilon_{\mathrm{b}K} = 2\gamma\mathcal{B} + J(2 + \lambda_{\mathrm{b}K}) < 2\gamma\mathcal{B} + J(2 - 2\cos\frac{Ka}{2}) \tag{30.232}$$

$$\varepsilon_K(q) \simeq 2\gamma\mathcal{B} + J(2 - 2\cos\frac{Ka}{2}\cos qa) \ , \quad -\frac{\pi}{a} \leq q \leq +\frac{\pi}{a} \tag{30.233}$$

Pour N fini, et pour tout K fixé, il y a donc un état discret (isolé) situé sous un ensemble d'états très proches en énergie les uns des autres. Dans la suite, on s'intéresse principalement à la solution $\lambda_{\mathrm{b}K}$, une question importante étant de savoir si, dans la limite $N \to +\infty$, elle se "colle" au bas du *continuum* ou si, au contraire, elle reste strictement au-dessous. Dans ce dernier cas, qui va de fait ressortir, il s'agit bel et bien d'une paire *liée* robuste, dont l'existence résiste à la limite $N \to +\infty$.

[80] Se souvenir que $K \in [-\frac{\pi}{a}, +\frac{\pi}{a}]$, de sorte que $\cos\frac{Ka}{2} \geq 0$.

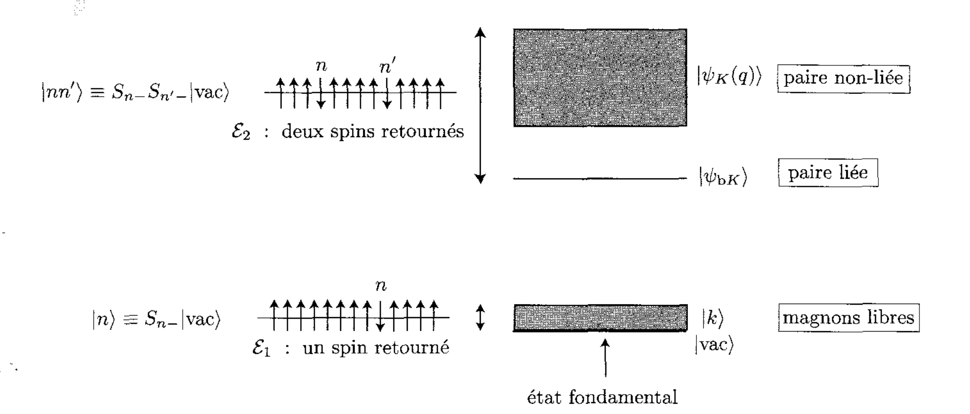


Figure 30.14: Schéma des excitations de basse énergie du modèle de Heisenberg ferromagnétique, $D = 1$. Dans l'état fondamental, tous les spins pointent en haut.

Pour savoir ce qu'il en est, il suffit de revenir à (30.231), en supposant justement $\lambda < -2\cos\frac{Ka}{2}$, puisque l'on s'intéresse à cette solution particulière. Avec cette restriction, le dénominateur de la fraction au second membre de (30.231) ne peut pas s'annuler, et il n'y a aucun risque à remplacer, quand $N \gg 1$, la somme par l'intégrale de Riemann dont elle est le précurseur. Ceci fait, l'équation aux valeurs prend la forme plus agréable :

$$1 = \frac{a}{\pi}\int_{-\pi/a}^{+\pi/a} \cos qa \, \frac{\cos\frac{Ka}{2} - \cos qa}{\lambda + 2\cos qa \cos\frac{Ka}{2}}\,\mathrm{d}q \ . \tag{30.234}$$

En posant $\lambda = -2\cosh u \cos\frac{Ka}{2}$ $(u \in \mathbb{R}_+)$, cette équation s'écrit :

$$\cos\frac{Ka}{2} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \cos x \, \frac{\cos x - \cos\frac{Ka}{2}}{\cosh u - \cos x}\,\mathrm{d}x \ ; \tag{30.235}$$

utilisant maintenant les égalités :

$$\sum_{n\in\mathbb{Z}} \mathrm{e}^{inx}\,\mathrm{e}^{-|n|u} = \frac{\sinh u}{\cosh u - \cos x} \iff \frac{1}{2\pi}\int_{-\pi}^{+\pi} \frac{\mathrm{e}^{-inx}}{\cosh u - \cos x}\,\mathrm{d}x = \frac{\mathrm{e}^{-|n|u}}{\sinh u} \ , \tag{30.236}$$

on trouve que la solution de (30.235) est $\mathrm{e}^{-u} = \cos\frac{Ka}{2}$, soit $\lambda = -(1+\cos^2\frac{Ka}{2})$. Revenant à (30.224), on obtient l'expression de la valeur propre dans la limite $N \gg 1$:

$$\boxed{E_{\mathrm{b}K} = E_{\mathrm{vac}} + 2\gamma\mathcal{B} + J\sin^2\frac{Ka}{2}} \tag{30.237}$$

L'énergie de liaison de la paire est donc $J(2 - 2\cos\frac{Ka}{2}) - J\sin^2\frac{Ka}{2}$, soit :

$$\boxed{E_{\text{liaison}}(K) = 4J\sin^4\frac{Ka}{4}} \tag{30.238}$$

La figure 30.14 schématise les excitations de basse énergie. L'ensemble des états dans $\mathcal{E}_2$ fait penser à un atome, avec un état discret (lié) et des états de diffusion (non-liés). La figure 30.15 donne une information plus précise en ce qui concerne les relations de dispersion pour le même sous-espace.

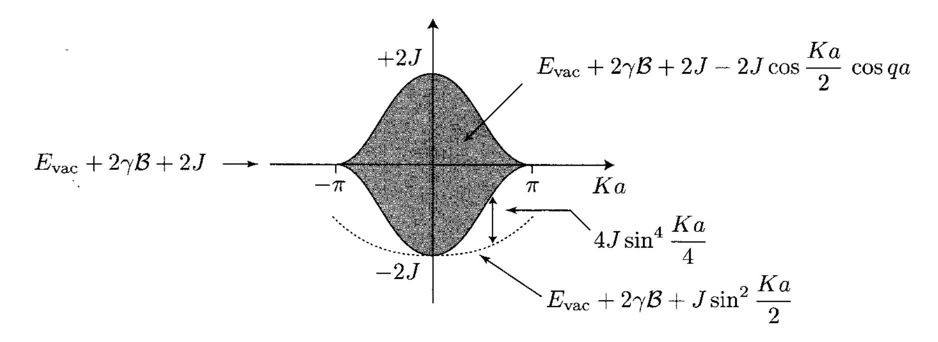


Figure 30.15: Lois de dispersion en fonction de K des excitations dans $\mathcal{E}_2$ du modèle de Heisenberg ferromagnétique, $D = 1$. La courbe en pointillés est l'état lié de la paire, dont l'énergie de liaison augmente avec K (voir (30.238)) ; la poche grisée représente le *continuum* des états non-liés, correspondant, à K fixé, à toutes les valeurs de q comprises entre $\pm\frac{\pi}{a}$.

Compte tenu de son caractère discret, l'état de la paire est un état lié, *molécule* de deux magnons. Ceci se confirme en analysant de plus près la fonction d'onde correspondante, c'est-à-dire les coefficients $c_{nn'}$, dont la variation en fonction de la distance $|n - n'|a$ constitue la distribution de probabilité spatiale des deux spins retournés. Les $c_{nn'}$ sont donnés par l'égalité (30.227), où la fonction $\Phi_K(q)$ est celle écrite en (30.230) avec $\lambda = -(1 + \cos^2\frac{Ka}{2})$:

$$\Phi_K(q) = \frac{\cos\frac{Ka}{2} - \cos qa}{-(1 + \cos^2\frac{Ka}{2}) + 2\cos qa\cos\frac{Ka}{2}}\, S_K \ , \qquad (30.239)$$

S_K étant une constante multiplicative que l'on peut avoir par normalisation. En remplaçant la somme par l'intégrale[81], et après normalisation, on trouve :

$$\boxed{c_{nn'} = \frac{1}{\sqrt{2}}\, \mathrm{e}^{\mathrm{i}\frac{K}{2}(n+n')a} \sin\frac{Ka}{2}\Big(\cos\frac{Ka}{2}\Big)^{|n-n'|-1} \qquad (|n-n'| \geq 1)} \qquad (30.240)$$

et avec toujours $c_{nn} = 0$. La fonction décroît exponentiellement avec la distance entre les deux spins retournés :

$$|c_{nn'}|^2 = \frac{1}{2}\sin^2\frac{Ka}{2}\, \mathrm{e}^{-2(|n-n'|-1)a/\xi} \ , \qquad \xi = \frac{a}{-\ln\big(\cos\frac{Ka}{2}\big)} \ . \qquad (30.241)$$

[81] Utiliser à nouveau (30.236).

ξ caractérise l'échelle de longueur de décroissance de la fonction d'onde de la paire liée :

$$\xi(K) \simeq \begin{cases} \frac{8}{K^2a^2} & K \simeq 0 \\ -\frac{1}{\ln(\pi - Ka)} & Ka \lesssim \pi \end{cases} . \tag{30.242}$$

La distance moyenne entre les deux spins basculés est $\langle d \rangle = \frac{a}{\sin^2(Ka/2)}$, sa fluctuation

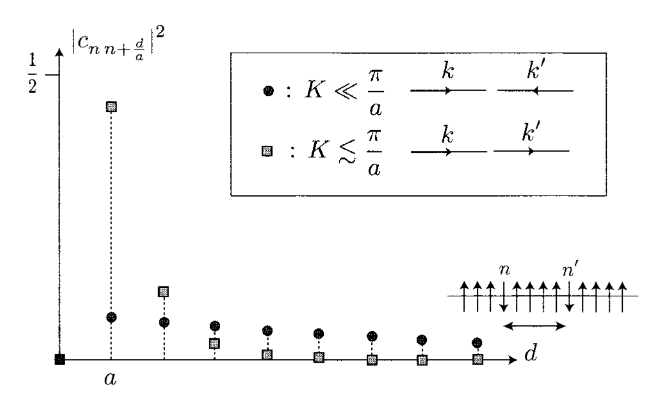


Figure 30.16: Fonction d'onde de la paire liée, suivant deux valeurs de l'impulsion totale $K \stackrel{\text{déf}}{=} \frac{1}{2}(k + k')$.

relative $\frac{\Delta d}{\langle d \rangle}$ vaut $\cos \frac{Ka}{2}$. C'est quand les deux magnons ont à peu près leur vitesse maximum, et vont dans le même sens, que la paire liée est très confinée, donnant la plus grande valeur pour l'énergie de liaison (voir (30.238)) ; alors, la probabilité de trouver deux spins basculés côte-à-côte domine toutes les autres. Au contraire, quand l'impulsion totale est très petite, en particulier si les impulsions sont presque opposées, la fonction d'onde est très étendue à l'échelle du pas du réseau.

L'existence de la paire liée est le résultat de plusieurs tendances antagonistes. Les deux spins renversés ont une répulsion de cœur dur puisqu'ils ne peuvent se trouver sur le même site ; d'un autre côté, ils sont en interaction attractive s'ils sont sur deux sites premiers voisins. Enfin, lorsqu'ils sont plus distants, la présence des autres spins tend à les faire revenir dans la position ferromagnétique. Le résultat net est une interaction effective attractive permettant la formation de la paire liée. L'augmentation de l'énergie de liaison avec l'impulsion totale évoque un écrantage de l'interaction effective d'autant plus faible que les deux magnons vont vite, et dans la même direction.

♦ *Remarque*

On a vu à plusieurs reprises le rôle important de la valeur $S = 1/2$ des spins pour la résolution exacte du problème. Pour un spin S quelconque, on dispose de schémas approximatifs utilisant une transformation dite *de Holstein - Primakoff* et les états cohérents de spin, permettant de définir une théorie *linéarisée* où les magnons constituent un gaz dilué de bosons. ♦

30.7 Exercices et problèmes

30.7.1 Modification de la sphère de Fermi pour un alcalin

Les alcalins (un électron de conduction par atome) cristallisent dans le système cubique centré (CC). Dans toute la suite, N est le nombre d'électrons quasi-libres, V est le volume de l'échantillon sur lequel sont définies les conditions cycliques de Born - von Kármán ; par ailleurs, on se place à température nulle.

Préliminaires

1. Soit k_{F}^0 le module du vecteur d'onde de Fermi calculé pour des électrons libres. Donner l'expression de k_{F}^0 en fonction de N et V. Quelle est l'énergie de Fermi $\varepsilon_{\mathrm{F}}^0$ correspondante ?
2. Préciser le réseau réciproque, $\tilde{\mathcal{B}}$. On rappelle que si a désigne le côté de la maille cubique conventionnelle du réseau direct $\mathcal{B}$, la longueur du côté de la maille conventionnelle de $\tilde{\mathcal{B}}$ est égale à $4\pi/a$.
3. Soit Γ le centre de la maille de Wigner - Seitz de $\tilde{\mathcal{B}}$ (première zone de Brillouin) et N l'un des douze plus proches voisins de Γ. Calculer la longueur ΓN. Dans toute la suite, on désigne par $\vec{K}$ le vecteur $\overrightarrow{\Gamma\mathrm{N}}$.
4. En déduire que la sphère de Fermi des électrons libres ne coupe pas le plan de Bragg associé à $\vec{K}$. Illustrer ceci par un schéma aussi explicite que possible.

Modification de la sphère libre

En raison du potentiel périodique (supposé faible), la loi de dispersion libre $\varepsilon^0(\vec{q})$:

$$\varepsilon^0(\vec{k}\,) = \frac{\hbar^2\vec{k}^{\,2}}{2m} \tag{30.243}$$

est modifiée près des plans de Bragg. Au voisinage du seul plan de Bragg de $\vec{K}$, l'approximation des électrons presque libres donne les deux bandes :

$$\varepsilon_{\pm}(\vec{k}\,) = \Sigma(\vec{k}\,) \pm \left[[\Delta(\vec{k}\,)]^2 + |\mathcal{V}(\vec{K})|^2\right]^{1/2} , \tag{30.244}$$

où :

$$\Sigma(\vec{k}\,) = \frac{1}{2}\,[\varepsilon^0(\vec{k}) + \varepsilon^0(\vec{K} - \vec{k}\,)] , \qquad \Delta(\vec{k}\,) = \frac{1}{2}\,[\varepsilon^0(\vec{k}) - \varepsilon^0(\vec{K} - \vec{k}\,)] , \tag{30.245}$$

et où $\mathcal{V}(\vec{K})$ est la composante de Fourier pertinente du potentiel de réseau. Dans la suite, on calculera en coordonnées sphériques (angles usuels θ et ϕ) en prenant l'axe Oz dirigé suivant $\vec{K}$.

1. Calculer $\Sigma(\vec{k})$ et $\Delta(\vec{k})$ en fonction de[82] k, K et θ.

2. S'agissant de trouver la surface de Fermi modifiée par le petit potentiel périodique, quel signe doit-on choisir dans l'expression (30.244) ? Dans la suite, on désigne simplement par $\varepsilon(\vec{k})$ la branche ainsi retenue et on suppose que $k < \frac{K}{2}$.

3. La surface d'énergie constante donnée $\varepsilon_{\rm F}$ a formellement pour équation :

$$\varepsilon(\vec{k}) = \varepsilon_{\rm F} \ . \qquad (30.246)$$

Expliciter cette relation en retenant les termes d'ordre 2 en $|\mathcal{V}_{\vec{K}}|$ au plus. Montrer qu'elle peut s'écrire :

$$k = \sqrt{\frac{2m\varepsilon_{\rm F}}{\hbar^2}}\left[1 + f(\theta)\frac{m|\mathcal{V}(\vec{K})|^2}{\hbar^2 K^2 \varepsilon_{\rm F}^0}\right] \qquad (30.247)$$

où f est une fonction sans dimension de la forme :

$$f(\theta) = \frac{1}{1 - C\,\cos\theta} \ , \qquad C = 2\frac{\sqrt{2m\varepsilon_{\rm F}^0}}{\hbar K} < 1 \ . \qquad (30.248)$$

Décrire qualitativement les modifications par rapport à la sphère libre de Fermi.

4. Le niveau de Fermi, $\varepsilon_{\rm F}$, s'obtient par l'équation :

$$2\int \frac{{\rm d}^3 k}{(2\pi)^3}\,\Theta\big(\varepsilon_{\rm F} - \varepsilon(\vec{k})\big) = \frac{N}{V} \ , \qquad (30.249)$$

où $\Theta(x)$ est la fonction échelon-unité. D'où vient le facteur 2 au premier membre ?

5. La contrainte exprimée par la fonction de Heaviside signifie que les niveaux sont remplis en k jusqu'à une certaine valeur $k_{\rm max}$ (dépendant de l'angle θ). Montrer que, toujours au même ordre d'approximation et lorsqu'un seul plan de Bragg est en jeu, le niveau de Fermi est calculable à partir de l'équation :

$$\frac{1}{3\pi^2}\left(\frac{2m\varepsilon_{\rm F}}{\hbar^2}\right)^{3/2}\left[1 + \frac{3m|\mathcal{V}(\vec{K})|^2}{2\hbar^2 K^2\varepsilon_{\rm F}^0}\,I\right] = \frac{N}{V} \ , \qquad (30.250)$$

où I désigne l'intégrale :

$$I \stackrel{\text{déf}}{=} \int_0^\pi \frac{\sin\theta}{1 - C\cos\theta}\,{\rm d}\theta \ . \qquad (30.251)$$

6. Déduire de ce qui précède que la variation de $\varepsilon_{\rm F}$ est donnée par :

$$\varepsilon_{\rm F} = \varepsilon_{\rm F}^0 - \frac{m|\mathcal{V}(\vec{K})|^2}{\hbar^2 K^2}\,I \ . \qquad (30.252)$$

[82] k désigne le module de $\vec{k}$, etc.

Calculer l'intégrale I pour obtenir l'expression finale de la variation du niveau de Fermi incluant au second ordre l'effet du potentiel de réseau pour un seul plan de Bragg :

$$\varepsilon_{\mathrm{F}} - \varepsilon_{\mathrm{F}}^{0} = -\frac{|\mathcal{V}(\vec{K})|^2}{4\varepsilon_{\mathrm{F}}^{0}}\frac{k_{\mathrm{F}}^{0}}{K}\ln\frac{K+2k_{\mathrm{F}}^{0}}{K-2k_{\mathrm{F}}^{0}} \tag{30.253}$$

Compte tenu du nombre de plans délimitant la première zone de Brillouin, quelle est la variation totale du niveau de Fermi ?

30.7.2 Modulation de l'énergie de site pour un réseau en liaisons fortes

On considère un réseau unidimensionnel de paramètre de maille a, traité en liaisons fortes, où l'énergie atomique varie d'un site à l'autre. Le Hamiltonien d'un électron se déplaçant sur le réseau est écrit sous la forme $H = H_0 + V$ avec :

$$H_0 = -t\sum_n (|n+1\rangle\langle n| + |n\rangle\langle n+1|) \quad (t>0) \ , \quad V = \sum_n v_n |n\rangle\langle n| \ . \tag{30.254}$$

Les états $\{|n\rangle\}_n$ sont des états atomiques se déduisant tous les uns des autres par translation ; $|n\rangle$ est l'état localisé autour du site d'abscisse $x = na$; les états $\{|n\rangle\}_n$ sont supposés orthonormalisés ($\langle n|n'\rangle = \delta_{nn'}$). Le cas échéant, on prendra en compte des conditions cycliques avec N sites.

La modulation de l'énergie de site est une onde plane d'amplitude v_0 et de nombre d'onde κ :

$$v(x) = v_0 \cos\kappa x \ , \qquad \kappa = \frac{2\pi}{\nu a} \ . \tag{30.255}$$

où ν et v_0 sont des quantités positives. v_n désigne la valeur de $v(x)$ au point $x = na$.

États de Bloch

1. Quelle différence qualitative existe-t-il entre le cas $\nu \in \mathbb{N}^*$ et ν non-entier ?
2. Dans le cas où $\nu = p \in \mathbb{N}^*$, dessiner une maille élémentaire du réseau.
3. Dans cette question, on ne considère que H_0 et on introduit les N états :

$$|\psi_k\rangle = \frac{1}{\sqrt{N}}\sum_n \mathrm{e}^{\mathrm{i}nka}|n\rangle \ , \qquad -\frac{\pi}{a} < k < +\frac{\pi}{a} \ . \tag{30.256}$$

   (a) Soit $\mathcal{T}_m$ l'opérateur associé à la translation ma. Combien vaut $\mathcal{T}_m|n\rangle$? Trouver $\mathcal{T}_m|\psi_k\rangle$.

   (b) En déduire que $|\psi_k\rangle$ est propre de H_0 et donner l'expression de la valeur propre correspondante, $\varepsilon^{(0)}(k)$.

Cas particulier $p = 2$

Dans cette partie, on choisit $p = 2$.

1. Quelle est une maille primitive ?
2. On pose :

$$|\psi_{k,0}\rangle = \frac{1}{\sqrt{N}} \sum_n \mathrm{e}^{inka} |n\rangle \ , \qquad |\psi_{k,1}\rangle = \frac{1}{\sqrt{N}} \sum_n (-1)^n \mathrm{e}^{inka} |n\rangle \qquad (30.257)$$

   où k est dans l'intervalle $[-\pi/(2a), +\pi/(2a)]$. Calculer $\mathcal{T}_m|\psi_{k,r}\rangle$. Qu'en déduit-on pour l'élément de matrice $\langle\psi_{k,r}|H|\psi_{k',r'}\rangle$ quand $k \neq k'$?

3. (a) Utiliser les résultats obtenus précédemment pour écrire $H_0|\psi_{k,r}\rangle$ ($r = 0, 1$).
   (b) Calculer $V|\psi_{k,r}\rangle$ ($r = 0, 1$).
4. En déduire la matrice 2×2 représentant H dans le sous-espace $|\psi_{k,r}\rangle$, $r = 0, 1$.
5. Trouver les valeurs propres de H, $\varepsilon_r(k)$, avec $\varepsilon_0(k) < \varepsilon_1(k)$. En déduire les lois de dispersion dans l'intervalle $[-\frac{\pi}{2a}, +\frac{\pi}{2a}]$ et les représenter graphiquement.
6. L'état propre associé à $\varepsilon_0(k)$ est noté :

$$|\Phi_{k,0}\rangle = \cos\theta_k|\psi_{k,0}\rangle - \sin\theta_k|\psi_{k,1}\rangle \qquad (30.258)$$

   Donner l'expression de $\tan\theta_k$.
7. En déduire l'expression de l'état propre $|\Phi_{k,1}\rangle$ associé à $\varepsilon_1(k)$.

Cas p entier et $p \geq 3$

Dans la suite, p est un entier positif ≥ 3. On définit les p vecteurs :

$$|\psi_{k,r}\rangle \stackrel{\text{déf}}{=} \frac{1}{\sqrt{N}} \sum_n \alpha_r^n \, \mathrm{e}^{inka} |n\rangle \ , \qquad -\frac{\pi}{pa} < k < +\frac{\pi}{pa} \qquad (30.259)$$

où les α_r sont les racines p^{es} de l'unité :

$$\alpha_r = \mathrm{e}^{ir\frac{2\pi}{p}} \qquad (r = 0, 1, 2, \ldots, p-1) \ . \qquad (30.260)$$

1. À quoi est égal $H_0|\psi_{k,r}\rangle$? En déduire les lois de dispersion dans la zone réduite $[-\frac{\pi}{pa}, +\frac{\pi}{pa}]$.
2. En calculant $V|\psi_{k,r}\rangle$, montrer que le sous-espace engendré par les p vecteurs $|\psi_{k,r}\rangle$ est stable par H.
3. Combien y a-t-il de valeurs propres $\varepsilon_r(k)$?
4. Sans faire de calcul, tracer qualitativement pour $p = 3$ le graphe des valeurs propres de H en fonction de k pour $k \in [-\frac{\pi}{pa}, +\frac{\pi}{pa}]$.

Remplissage fractionnaire particulier

On suppose que dans le cristal une fois constitué, il y a en moyenne 2/3 d'électron par atome.

1. Pour le cristal non modulé ($\kappa = 0$), quel est le taux de remplissage de la bande ?
2. Le solide est-il conducteur ou isolant ?
3. Indiquer un moyen permettant d'inverser la propriété de conduction du solide.

30.7.3 Règle de sélection pour un cristal parfait

Il s'agit d'établir la règle de sélection impliquant le vecteur d'onde $\vec{k}$ caractérisant les états propres électroniques d'un cristal supposé sans vibrations, et d'en donner une interprétation physique. Dans toute la suite, $\vec{k}$ est choisi dans la première zone de Brillouin.

Définition du modèle

1. Les seuls états excités considérés dans la suite sont des états de Bloch $|\psi_{\vec{k}}\rangle$; écrire la relation entre $\psi_{\vec{k}}(\vec{r}+\vec{R})$ et $\psi_{\vec{k}}(\vec{r})$.
2. L'état fondamental du cristal, noté $|\psi_{\rm g}\rangle$, non-dégénéré, est invariant dans toute translation $\mathcal{T}(\vec{R})$ du réseau de Bravais $\mathcal{B}$:
$$\psi_{\rm g}(\vec{r}+\vec{R}) = \psi_{\rm g}(\vec{r}) \ . \tag{30.261}$$
Sachant que cet état, lui aussi, satisfait le théorème de Bloch, quelle est la valeur du vecteur d'onde de cet état ?

Perturbation du cristal par un champ électromagnétique

Le cristal est soumis à une onde plane monochromatique de vecteur d'onde $\vec{q}$, traitée semi-classiquement. Dans ces conditions, l'amplitude de transition $A_{g\rightarrow\vec{k}}$ de l'état fondamental à l'état excité $|\psi_{\vec{k}}\rangle$ est (au premier ordre) proportionnelle à une intégrale (temporelle) contenant l'élément de matrice $\langle\psi_{\vec{k}}|{\rm e}^{{\rm i}\vec{q}.\vec{r}}|\psi_{\rm g}\rangle$.

1. Pour un cristal tridimensionnel, on a :
$$\langle\psi_{\vec{k}}|{\rm e}^{{\rm i}\vec{q}.\vec{r}}|\psi_{\rm g}\rangle = \int \psi^*_{\vec{k}}(\vec{r}){\rm e}^{{\rm i}\vec{q}.\vec{r}}\psi_{\rm g}(\vec{r})\,{\rm d}^3r \ . \tag{30.262}$$
En effectuant le changement de variable $\vec{r} = \vec{r}\,' + \vec{R}$, établir la relation entre les deux intégrales $\int {\rm d}^3r' \ldots$ et $\int {\rm d}^3r \ldots$

2. À quelle condition l'amplitude de transition est-elle forcément nulle à cet ordre ?

3. *A contrario*, exprimer d'une phrase impliquant le réseau réciproque $\tilde{\mathcal{B}}$ la condition sur $\vec{k} - \vec{q}$ pour que la transition soit possible[83]. En donner une interprétation physique.

4. Quelle autre condition le résultat évoque-t-il ?

5. On suppose que la différence d'énergie entre $|\psi_{\mathrm{g}}\rangle$ et les $|\psi_{\vec{k}}\rangle$ est de l'ordre de 2 eV. Quel est, pour une onde résonnante, l'ordre de grandeur de la longueur d'onde λ et du module du vecteur d'onde $\vec{q}$?

6. Justifier l'appellation de *transition verticale* pour une transition dont le vecteur $\vec{K} \in \tilde{\mathcal{B}}$ apparaissant dans la condition 3 est égal à zéro (transition dite *normale*).

7. Pour une transition verticale à partir de l'état fondamental, quelle est la valeur de la quasi-impulsion de l'état excité accessible ?

8. Sans calcul, mais en donnant une explication claire et précise, écrire et interpréter la condition exprimant que la transition *normale* intrabande $|\psi_{\vec{k}}\rangle \longrightarrow |\psi_{\vec{k}'}\rangle$ est permise (l'illustrer par un petit dessin *à la* Feynman)

Extension au cas d'un cristal bi- ou uni-dimensionnel

1. Pour un réseau bi-dimensionnel $\mathbb{Z}^2$, de vecteurs primitifs $\vec{a}_1$ et $\vec{a}_2$, quelle est la forme du vecteur $\vec{R}$?

2. On introduit transitoirement un vecteur $\vec{a}_3$ perpendiculaire aux deux vecteurs $\vec{a}_i$ de $\mathbb{Z}^2$. On désigne par $\tilde{\mathcal{B}}$ le réseau réciproque construit sur trois vecteurs primitifs $\vec{b}_j$ ($i = j$, 2, 3) formés en partant des relations usuelles entre les $\vec{b}_j$ et les $\vec{a}_i$. Rappeler ces relations.

3. On note θ l'angle entre $\vec{a}_1$ et $\vec{a}_2$. Calculer les modules des $\vec{b}_j$.

4. Pour engendrer de fait un réseau bidimensionnel, on fait tendre le module de $\vec{a}_3$ vers l'infini. Que valent les modules des $\vec{b}_j$ à la limite ?

5. Suivant ce procédé, quelle est la dimension du réseau réciproque du réseau bidimensionnel ?

6. Quelle que soit la dimensionnalité du cristal, on a trouvé précédemment la condition :
$$\mathrm{e}^{\mathrm{i}(\vec{q}-\vec{k})\vec{R}} = 1 \tag{30.263}$$
exprimant que la transition $|\psi_{\mathrm{g}}\rangle \longrightarrow |\psi_{\vec{k}}\rangle$ est permise. En déduire la forme précise de $\vec{k} - \vec{q}$ pour que cette transition soit permise pour un réseau bidimensionnel. Illustrer par un schéma.

[83] En pareil cas, on dira simplement dans la suite que la transition est *permise*.

7. En se bornant aux transitions *normales*, quelle peut être la quasi-impulsion $\vec{k}$ des états excités accessibles par perturbation du champ électromagnétique ?

8. Expliquer brièvement comment ces résultats se transforment dans le cas d'un réseau uni-dimensionnel.

30.7.4 Un modèle pour le graphite

La variété graphite du carbone est constituée d'un empilement de plans parallèles contenant chacun un réseau en nid d'abeille. La relativement grande distance entre deux plans consécutifs[84] justifie que l'on puisse, pour certaines propriétés, considérer cette structure comme un empilement de plans sans interaction mutuelle d'aucune sorte. Cette hypothèse étant admise, l'étude du graphite se réduit à la considération d'un seul plan, l'objectif étant ici d'analyser la conductivité bidimensionnelle.

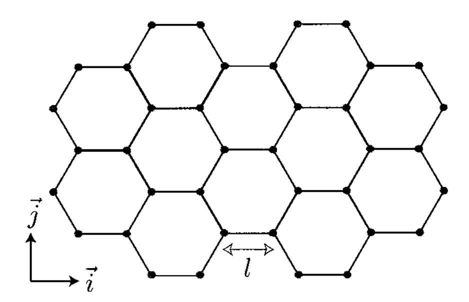


Figure 30.17: Fragment du réseau plan en nid d'abeille

Dans tout le problème, on considère donc un réseau plan en nid d'abeille, $\mathcal{R}$, construit sur des hexagones de côté l dont chaque sommet est occupé par un atome de carbone. Le cas échéant, on rapportera ce plan $\mathbb{R}^2$ à un repère orthonormé $(\vec{i}, \vec{j})$ (voir fig. 30.17).

1. **Étude de la structure cristalline**

   (a) Le réseau nid d'abeille est-il un réseau de Bravais ?

   (b) Préciser un réseau de Bravais $\mathcal{B}$ et son motif (base) permettant d'engendrer le réseau $\mathcal{R}$.

   (c) Soit $\vec{a}_1$, $\vec{a}_2$ deux vecteurs primitifs de $\mathcal{B}$ formant un angle aigu, $\vec{a}_1$ étant choisi le long de $\vec{i}$. Exprimer leur longueur a en fonction de l. Quelles sont les composantes de $\vec{a}_2$?

   (d) Quel est le "volume" d'une maille primitive de $\mathcal{B}$?

[84]La distance entre deux plans consécutifs est environ 2,4 fois supérieure à la distance entre deux atomes voisins d'un même plan.

(e) Trouver les deux vecteurs primitifs $\vec{b}_1$, $\vec{b}_2$ du réseau réciproque associés à $\vec{a}_1$, $\vec{a}_2$ (les exprimer sur la base $(\vec{i}, \vec{j})$). Quel est le "volume" d'une maille primitive de $\tilde{\mathcal{B}}$?

(f) Indépendamment du motif, à quelle condition sur le transfert de moment $\vec{k}_\text{f} - \vec{k}_\text{i}$ observe-t-on une tache de diffraction ?

(g) Soit $F(\vec{q})$ le facteur de structure géométrique associé au motif. En désignant par $\vec{R}$ et $\vec{R} + \vec{d}$ les rayons-vecteurs de deux éléments du motif, écrire l'expression de $F(\vec{q})$.

(h) Trouver la(es) condition(s) d'extinction des raies compte tenu du motif.

2. **Structure électronique dans l'approximation des liaisons fortes**

Dans toute la suite, on définit une fonction atomique de type[85] 2p_0 sur chaque site de $\mathcal{R}$; la fonction centrée sur le site en $\vec{R}$ est notée $\phi_{\vec{R}}$, celle centrée sur $\vec{R} + \vec{d}$ est notée de même $\phi_{\vec{R}+\vec{d}}$. On négligera tous les produits scalaires entre deux fonctions atomiques localisées sur deux sites différents. Le cas échéant, on invoquera des conditions cycliques de Born - von Kármán définies sur N mailles primitives de $\mathcal{B}$.

Soit H_0 le Hamiltonien d'un électron du motif isolé localisé en $\vec{R}$. On pose :

$$\langle\phi_{\vec{R}}|H_0|\phi_{\vec{R}}\rangle = \langle\phi_{\vec{R}+\vec{d}}|H_0|\phi_{\vec{R}+\vec{d}}\rangle = \varepsilon_0 \ , \qquad \langle\phi_{\vec{R}}|H_0|\phi_{\vec{R}+\vec{d}}\rangle = v \qquad (30.264)$$

(a) Pourquoi les éléments de matrice (30.264) sont-ils indépendants[86] de $\vec{R}$?

(b) v est réel : pourquoi ? Dans la suite, on suppose les phases choisies de sorte que v soit positif.

(c) Quelles sont les valeurs propres $E_\pm$ ($E_+ > E_-$) de H_0 ? Écrire les vecteurs propres correspondants $|\phi_{\vec{R}\,\pm}\rangle$.

(d) Utiliser la base des $|\phi_{\vec{R}\,\pm}\rangle$ pour former deux vecteurs $|\psi_{\vec{k}\,\pm}\rangle$ satisfaisant le théorème de Bloch.

(e) Tracer la première zone de Brillouin.

(f) Normaliser les $|\psi_{\vec{k}\,\pm}\rangle$.

(g) Soit H le Hamiltonien d'un électron sur le réseau ; combien vaut l'élément de matrice $\langle\psi_{\vec{k}\,\eta}|H|\psi_{\vec{k}'\,\eta'}\rangle$, $\forall\eta$, $\eta' = \pm$, si $\vec{k} \neq \vec{k}'$? Expliquer pourquoi il en est ainsi.

(h) Montrer que :

$$\langle\psi_{\vec{k}\eta}|H|\psi_{\vec{k}\eta'}\rangle = \sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}\,\langle\phi_{\vec{0}\,\eta}|H|\phi_{\vec{R}\,\eta'}\rangle \ . \qquad (30.265)$$

(i) On écrit $H = H_0 + \Delta V$ où ΔV est l'énergie potentielle due aux atomes autres que ceux du motif en $\vec{R} = 0$. En se limitant aux termes venant des atomes seconds voisins au plus, on pose :

$$\langle\phi_{\vec{0}}|\Delta V|\phi_{\vec{a}_i}\rangle = w \qquad (w > 0) \ . \qquad (30.266)$$

---

[85] Cette notation désigne une fonction construite avec l'harmonique sphérique $Y_{l=1\,m=0}$.

[86] Ceci étant, on pourra considérer dans la suite que H_0 est relatif au motif situé en $\vec{R} = \vec{0}$.

Avec ces notations, montrer que la matrice de H sur la base $|\psi_{\vec{k}\,+}\rangle$, $|\psi_{\vec{k}\,-}\rangle$ est donnée par[87] :

$$H = \begin{bmatrix} v(1+f) + 2wg & ivh \\ -ivh & -v(1+f) + 2wg \end{bmatrix} , \qquad (30.267)$$

où f, g et h sont des fonctions, symétriques et à valeurs réelles, des deux produits scalaires $\vec{k}.\vec{a}_1$ et $\vec{k}.\vec{a}_2$; montrer qu'elles satisfont la relation :

$$(1+f)^2 + h^2 = 3 + 2g \ .$$

Trouver l'expression des deux valeurs propres $\varepsilon_\pm(\vec{k})$ en fonction de v, w et de la seule fonction g.

3. **Discussion**

(a) Pour N mailles primitives de $\mathcal{B}$, combien existe-t-il d'états $\vec{k}$ distincts pour une bande d'énergie donnée ?

(b) Sachant que chaque atome de carbone apporte un électron, combien y a-t-il d'électron(s) par maille élémentaire ?

(c) Dans l'hypothèse où les deux bandes $\varepsilon_\pm$ sont disjointes, un plan de graphite est-il conducteur ou isolant à température nulle ?

(d) Montrer que la fonction g introduite ci-dessus a pour expression :

$$g(\vec{k}\,) = \cos\vec{k}.\vec{a}_1 + \cos\vec{k}.\vec{a}_2 + \cos\vec{k}.(\vec{a}_1 - \vec{a}_2) \qquad (30.268)$$

et satisfait $-\frac{3}{2} \leq g(\vec{k}\,) \leq 3 \ \forall\vec{k}\ \in$BZ1 ; la valeur minimale de g est atteinte aux sommets de BZ1, points notés Γ_i. Quelle est la relation entre les $\varepsilon_\pm(\vec{k}\,)$ aux points Γ_i ?

(e) Si $w = 0$, y a-t-il *recouvrement* entre les deux bandes $\varepsilon_\pm(\vec{k}\,)$?

(f) Qu'en concluez-vous pour la conductivité d'un plan de graphite quand $w = 0$?

(g) Les deux valeurs propres $\varepsilon_\pm$ sont égales respectivement à deux fonctions $\alpha_\pm(g)$. Tracer sommairement le graphe de celles-ci en fonction de g.

(h) On désigne par g_0 la valeur de g telle que $\varepsilon_-(g_0) = \varepsilon_+(-3/2)$. À quelle condition sur g_0 la bande $\varepsilon_+(\vec{k}\,)$ est-elle partiellement remplie ? Montrer que ceci se produit si le rapport $\frac{w}{v}$ est supérieur à une certaine valeur λ_0. Combien vaut λ_0 ?

(i) En déduire que si $w/v > \lambda_0$, le présent modèle prévoit que le graphite est un semi-métal.

[87] à la constante additive $\varepsilon_0\mathbf{1}$ près.

30.7.5 États d'un électron presque libre sur un réseau hexagonal bidimensionnel

On considère un réseau de Bravais monoatomique hexagonal $\mathcal{B}$ dont deux vecteurs primitifs sont $\vec{a}_1$ et $\vec{a}_2$, de même longueur a et faisant entre eux un angle de $2\pi/3$. On désigne par $V(\vec{r})$ le potentiel périodique du réseau et par v le "volume" de la maille primitive de $\mathcal{B}$. $\vec{k}$ désigne le pseudo-moment de l'électron dans le cristal.

1. (a) Décrire le mode de construction de la maille de Wigner - Seitz et la dessiner.

   (b) Soit l'exponentielle $\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}$ où $\vec{R} \in \mathcal{B}$; combien vaut-elle si $\vec{k} \in \tilde{\mathcal{B}}$, où $\tilde{\mathcal{B}}$ est le réseau réciproque plan ?

   (c) Soit $\vec{b}_1$ et $\vec{b}_2$ les deux vecteurs primitifs de $\tilde{\mathcal{B}}$ construits avec $\vec{a}_1$ et $\vec{a}_2$ et n'importe quel autre vecteur $\vec{a}_\perp$ orthogonal aux $\vec{a}_i$. Expliquer rapidement pourquoi $\vec{b}_1$ est donné par :

$$\vec{b}_1 = 2\pi \frac{\vec{a}_2 \times \vec{a}_\perp}{(\vec{a}_1, \vec{a}_2, \vec{a}_\perp)} \quad . \tag{30.269}$$

   (d) i. Calculer la longueur des vecteurs $\vec{b}_i$.

   ii. Dessiner la maille primitive de $\tilde{\mathcal{B}}$ construite sur ces vecteurs, ainsi que la première zone de Brillouin (BZ1).

   Δ_1 et Δ_2 étant les "plans bissecteurs" des deux vecteurs $\vec{b}_1$ et $\vec{b}_2$, on désigne dans la suite par M_i le point d'intersection de $\vec{b}_i$ avec Δ_i, par P le point d'intersection de Δ_1 et Δ_2, et par O le centre de BZ1.

   (e) Soit $\psi_{\vec{k}}(\vec{r})$ un état de Bloch ; rappeler la relation entre $\psi_{\vec{k}}(\vec{r})$ et $\psi_{\vec{k}}(\vec{r}+\vec{R})$, quel que soit $\vec{R} \in \mathcal{B}$.

   (f) Pourquoi la moyenne spatiale du potentiel de réseau peut-elle être prise égale à zéro ?

2. (a) On désigne par $\varepsilon^{(0)}(\vec{k})$ la relation de dispersion d'un électron libre, et on définit les trois fonctions :

$$\varepsilon_0^{(0)}(\vec{k}) \overset{\text{déf}}{=} \varepsilon^{(0)}(\vec{k}) \ , \qquad \varepsilon_i^{(0)}(\vec{k}) \overset{\text{déf}}{=} \varepsilon^{(0)}(\vec{k}-\vec{b}_i) \qquad (i=1,\,2) \ . \tag{30.270}$$

   Afin d'analyser la dépendance en $\vec{k}$ lorsque $\vec{k} \parallel \vec{b}_1$, on pose :

$$\vec{k} = \lambda \vec{b}_1 \qquad (\lambda \in \mathbb{R}) \ . \tag{30.271}$$

   i. Combien vaut λ en bord de zone ?

   ii. Tracer les graphes des deux énergies $\varepsilon_0^{(0)}(\vec{k})$ et $\varepsilon_1^{(0)}(\vec{k})$ quand le vecteur $\vec{k}$ varie de la sorte, tout en restant confiné dans BZ1.

   iii. Quelle est la relation entre $\varepsilon_0^{(0)}$ et $\varepsilon_1^{(0)}$ pour $\vec{k} = \vec{b}_1/2$?

(b) S'agissant de trouver les états propres approchés de l'électron en présence du potentiel $V(\vec{r})$ pour $\vec{k} \simeq \vec{b}_1/2$, rappeler pourquoi on peut se borner à considérer les deux types d'états d'ordre zéro associés aux valeurs propres $\varepsilon_0^{(0)}$ et $\varepsilon_1^{(0)}$.

(c) Soit les trois ondes planes :

$$\psi_0(\vec{r}) = \frac{1}{\sqrt{v}}\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}} \ , \qquad \psi_i(\vec{r}) = \frac{1}{\sqrt{v}}\,\mathrm{e}^{\mathrm{i}(\vec{k}-\vec{b}_i).\vec{r}} \qquad (i = 1,\, 2) \ . \tag{30.272}$$

Montrer que ces fonctions satisfont le théorème de Bloch.

(d) Lorsque $\vec{k}$ est au voisinage de M_1, on traite le problème par perturbation, en ne considérant que les deux états ψ_0 et ψ_1. Soit V_{01} l'élément de matrice :

$$V_{01} = \langle \psi_0 | V | \psi_1 \rangle \ . \tag{30.273}$$

Exprimer V_{01} en fonction de la transformée de Fourier $\mathcal{V}(\vec{q})$ définie comme :

$$\mathcal{V}(\vec{q}) = \frac{1}{v} \int_v \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{r}}\, V(\vec{r})\, \mathrm{d}^2 r \ . \tag{30.274}$$

$\mathcal{V}(\vec{q})$ est réel : pourquoi ?

(e) Écrire la matrice représentant le Hamiltonien (effectif) d'un électron dans le sous-espace engendré par ψ_0 et ψ_1 en fonction de $\vec{k}$, $\vec{b}_1$ et V_{01}.

(f) i. En déduire les deux valeurs propres approchées $\varepsilon_\pm(\vec{k})$.
  ii. Représenter leur variation en fonction de $\vec{k}$ donné par (30.271).
  iii. Exprimer le gap en bord de zone, E_g, en fonction de V_{01}.
  iv. Soit v_g la vitesse de groupe d'un électron.
    A. Quelle est son expression en fonction des $\varepsilon_\pm(\vec{k})$?
    B. Combien vaut-elle en bord de zone ?

3. Dans cette partie, on examine ce qui se passe quand l'extrémité de $\vec{k}$ est au *voisinage* du point P.

(a) Quelle est la relation entre les $\varepsilon_i^{(0)}$ $(i = 0, 1, 2)$ lorsque l'extrémité de $\vec{k}$ est en P ?

(b) Les nouvelles valeurs propres $\varepsilon(\vec{k})$ sont-elles extrémales en P le long de la direction de $\vec{b}_1 + \vec{b}_2$?

(c) Décrire la méthode de résolution à utiliser lorsque l'extrémité de $\vec{k}$ est proche du point P.

(d) Exprimer les éléments de matrice V_{01}, V_{02}, V_{12} à l'aide de $\mathcal{V}(\vec{q})$.

(e) En déduire l'allure précise de la matrice à diagonaliser pour décrire convenablement le voisinage du point P ?

(f) En formant des combinaisons linéaires simples des fonctions de base, montrer que le système 3×3 se factorise en deux blocs 1×1 et 2×2.

(g) En déduire les valeurs propres ; retrouver le comportement prévu en examinant le voisinage du point P.

30.7.6 États localisés dus à une impureté

On considère un réseau de Bravais $\mathcal{B} \equiv \{\vec{R}\}$. Sur chaque site est défini un état électronique bien localisé noté $|\vec{R}\rangle$, tous ces états étant orthonormalisés : $\langle\vec{R}|\vec{R}'\rangle = \delta_{\vec{R}\vec{R}'}$. Le cas échéant, on supposera adoptées des conditions cycliques, N désignant le nombre de nœuds du réseau.

1. Former les états $|\vec{k}\rangle$ satisfaisant le théorème de Bloch et les normaliser ; les énergies propres correspondantes sont notées $\varepsilon(\vec{k})$. Que peut-on dire de la densité électronique des états $|\vec{k}\rangle$, comparée à celle des états $|\vec{R}\rangle$?

2. On dope le réseau en introduisant une impureté au site $\vec{R}_0$. H_0 désignant le Hamiltonien du réseau non dopé, on a maintenant $H = H_0 + V$, avec $V = v|\vec{R}_0\rangle\langle\vec{R}_0|$ ($v \in \mathbb{R}$). Interpréter physiquement la perturbation due à l'impureté.

3. On introduit les opérateurs (résolvantes) $G(z)$ et $G_0(z)$ définis comme :

$$G(z) = \frac{1}{z\mathbf{1} - H} \ , \qquad G_0(z) = \frac{1}{z\mathbf{1} - H_0} \ . \tag{30.275}$$

   Démontrer que la relation suivante est vraie (équation de Dyson) :

$$G(z) = G_0(z) + G_0(z)VG(z) \ . \tag{30.276}$$

4. Quelle est l'expression de l'opérateur G_0 quand on le développe sur la base $\{|\vec{k}\rangle\}$?

5. La résolvante est une sorte de transformée de Fourier de l'opérateur avancé $U_+(t)$ défini en (23.130). Validez sommairement cette affirmation en considérant l'intégrale $\int_{\mathcal{C}_+} e^{\frac{1}{i\hbar}zt} G(z)\,dz$, où $\mathcal{C}_+$ est la droite parallèle à l'axe réel, située juste au-dessus de ce dernier. En définitive, les énergies propres, zéros de $\text{Dét}(z\mathbf{1} - H)$, ne sont rien d'autre que les pôles de la résolvante et donnent les fréquences de Bohr du système.

6. On pose :

$$g_0(z) = \langle\vec{R}_0|G_0|\vec{R}_0\rangle \ , \qquad g(z) = \langle\vec{R}_0|G|\vec{R}_0\rangle \ . \tag{30.277}$$

   Utiliser l'équation de Dyson pour trouver l'expression de $g(z)$ en fonction de $g_0(z)$ et v.

7. En déduire l'équation formelle donnant les énergies propres du réseau avec l'impureté

8. En utilisant le développement obtenu en 4, écrire g_0 sous la forme d'une somme de fractions rationnelles.

9. Expliciter l'équation écrite formellement en 7 sous la forme :

$$\text{quelque chose} = \frac{1}{v} \ . \tag{30.278}$$

10. Il s'agit ici de discuter graphiquement l'équation obtenue en 9 ; alors que N est en pratique $\sim 10^{23}$, on fera les dessins avec N de l'ordre de quelques unités (5 à 10).
   (a) Vérifier que l'équation (30.278) a bien exactement N solutions. Visiblement, l'une des racines se distingue nettement des autres : on l'appelle *racine singulière* dans la suite.
   (b) Montrer que quand $v \to 0$, on retrouve bien les énergies du réseau sans impureté.
   (c) Soit $[\varepsilon_{\text{min}}, \varepsilon_{\text{max}}]$ l'intervalle contenant toutes les énergies du réseau sans impureté. Comment se situe la racine singulière par rapport à cet intervalle selon le signe de v ?
   (d) Que peut-on dire des autres racines quand $N \gg 1$?
   (e) Quand v est négatif, et grand en module par rapport à toutes les autres énergies, l'impureté constitue un piège profond. À quoi ressemble dans ces conditions la fonction propre ψ_{sing} associée à la racine singulière ?
   (f) D'une façon générale (lorsque v est quelconque), comment se distingue ψ_{sing} des états étendus du réseau parfait ?
11. En présence de plusieurs bandes d'énergie, décrire physiquement l'effet du dopage. Quelles en sont les conséquences sur la conductivité électrique ?
12. Un semi-conducteur pur est-il conducteur à température nulle ? Expliquer sommairement pourquoi, toujours à température nulle, un semi-conducteur dopé a une conductivité finie.

30.7.7 Variation en température de la chaleur spécifique d'un solide possédant des excitations sans *gap*

On suppose que les premières excitations d'un solide en dimension D ont une relation de dispersion isotrope $\varepsilon(\vec{k}) = C\|\vec{k}\|^{\alpha}$, $\alpha > 0$; la répartition en température du nombre de ces excitations est une certaine fonction $f(u)$ où $u \stackrel{\text{déf}}{=} \beta\varepsilon$.

Montrer que l'énergie moyenne à la température T varie comme $T^{\frac{D}{\alpha}+1}$ et en déduire la variation de la chaleur spécifique c_V.

30.7.8 Magnétisme localisé : le modèle d'Anderson

Le modèle analysé ici a été proposé par Anderson [248] pour démontrer la possibilité de formation de moments magnétiques *localisés* résultant de la corrélation électronique quand on dissout une impureté dans un métal.

En schématisant à l'extrême, le métal est représenté par une seule bande d'énergie $E(\vec{k})$ associées aux états $|\vec{k}\rangle$:

$$H_{\mathrm{m}} = \sum_{\vec{k},s} E(\vec{k})\, c_{\vec{k},s}^{\dagger} c_{\vec{k},s} \qquad (s = \pm 1/2)\ . \tag{30.279}$$

La densité d'états $\rho(E)$ est supposée non nulle au-dessus d'une certaine valeur E_{min} et, à part cela, est quelconque ; on note ε_{F} le niveau de Fermi, $\varepsilon_{\mathrm{F}} > E_{\mathrm{min}}$.

L'impureté apporte un seul état non dégénéré, $|\mathrm{i}\rangle$, d'énergie E_{i}, pouvant recevoir un ou deux électrons de spin quelconque ; compte tenu du caractère localisé de cette orbitale, il importe de tenir compte de la répulsion entre électrons, ce que l'on fait à la manière de Hubbard en posant le Hamiltonien de l'impureté sous la forme :

$$H_{\mathrm{i}} = \sum_{s} E_{\mathrm{i}}\, c_{\mathrm{i},s}^{\dagger} c_{\mathrm{i},s} + \frac{U}{2} \sum_{s} n_s\, n_{-s}\ , \qquad n_s \stackrel{\text{déf}}{=} c_{\mathrm{i},s}^{\dagger} c_{\mathrm{i},s}\ . \tag{30.280}$$

Le paramètre $U > 0$ mesure la répulsion de deux électrons de spins contraires.

Il reste à traduire la possibilité pour un électron de passer de la bande d'états étendus à l'état localisé, ou inversement, ce que l'on fait en modélisant comme suit :

$$V = \sum_{\vec{k},s} V(\vec{k})\, c_{\vec{k},s}^{\dagger} c_{\mathrm{i},s} + V^*(\vec{k})\, c_{\mathrm{i},s}^{\dagger} c_{\vec{k},s}\ . \tag{30.281}$$

Le cas intéressant est celui où le niveau de Fermi est un peu au-dessus de l'énergie de l'impureté ; dans toute la suite on suppose donc $\varepsilon_{\mathrm{F}} \gtrsim E_{\mathrm{i}}$.

Calcul de la résolvante approchée

Ne sachant pas trouver les états propres exacts de $H \stackrel{\text{déf}}{=} H_{\mathrm{m}} + H_{\mathrm{i}} + V$, on adopte désormais une approximation de champ moyen consistant à remplacer $n_s n_{-s}$ par $\langle n_s\rangle n_{-s} + n_s \langle n_{-s}\rangle$, $\langle n_s\rangle$ désignant une valeur moyenne à calculer en temps utile par une équation auto-cohérente. Une fois effectuée cette *linéarisation*, on obtient un nouvel Hamiltonien H_{mf}.

1. Écrire le système d'équations pour les éléments de la matrice de la résolvante $G(z) \stackrel{\text{déf}}{=} (z\mathbf{1} - H_{\mathrm{mf}})^{-1}$ sur la base $(|\mathrm{i}, \mathrm{s}\rangle, |\vec{k}, s\rangle)$.

2. En déduire que l'élément diagonal $G_{\mathrm{i},s}(z) \stackrel{\text{déf}}{=} \langle \mathrm{i}, s|G(z)|\mathrm{i}, s\rangle$ est donné par :

$$G_{\mathrm{i},s}(z) = \frac{1}{z - E_s - \sum_{\vec{k}} \frac{|V(\vec{k})|^2}{z - E(\vec{k})}}\ . \tag{30.282}$$

   où on a posé $E_s \stackrel{\text{déf}}{=} E_{\mathrm{i}} + U\langle n_{-s}\rangle$.

3. En effectuant la substitution habituelle $\sum_{\vec{k}} \rightarrow \left(\frac{L}{2\pi}\right)^3 \int \mathrm{d}^3k$, exprimer $G_{\mathrm{i},s}(z)$ en fonction de la quantité $\Delta(E)$ définie comme :

$$\Delta(E) \stackrel{\text{déf}}{=} \frac{2\pi}{\hbar}\left(\frac{L}{2\pi}\right)^3 \int |V(\vec{k})|^2 \, \delta(E - E(\vec{k})) \, \mathrm{d}^3k \ . \tag{30.283}$$

En déduire la contribution $\rho_{\mathrm{i},s}(E)$ de l'état localisé à la densité d'états totale. Dans la suite, on néglige le déplacement (la *renormalisation*) de l'énergie de l'état localisé.

Équations auto-cohérentes

Le nombre moyen d'électrons dans l'état localisé est :

$$\langle n_s \rangle = \int_{E_{\mathrm{min}}}^{\varepsilon_{\mathrm{F}}} \rho_{\mathrm{i},s}(E) \, \mathrm{d}E \simeq \int_{-\infty}^{\varepsilon_{\mathrm{F}}} \rho_{\mathrm{i},s}(E) \, \mathrm{d}E \ . \tag{30.284}$$

1. On supose que la fonction $\Delta(E)$ est à variation lente, ce qui autorise à l'extraire de l'intégrale, en la remplaçant par sa valeur $\Delta(E_s) \stackrel{\text{déf}}{=} \Delta$. Montrer alors que la valeur moyenne $\langle n_s \rangle$ a pour expression :

$$\langle n_s \rangle = \frac{1}{2} + \frac{1}{\pi} \mathrm{Arctan} \frac{2(\varepsilon_{\mathrm{F}} - E_s)}{\hbar\Delta} \ . \tag{30.285}$$

2. On pose $\delta \stackrel{\text{déf}}{=} \frac{\varepsilon_{\mathrm{F}} - E_{\mathrm{i}}}{U}$, $\gamma \stackrel{\text{déf}}{=} \frac{\hbar\Delta}{2U}$. Écrire les deux équations couplées auto-cohérentes pour $\langle n_\uparrow \rangle$ et $\langle n_\downarrow \rangle$.
3. Construire graphiquement dans le plan $(\langle n_\uparrow \rangle, \langle n_\downarrow \rangle)$ le lieu des solutions de ce système dans l'hypothèse $\delta \sim 1/2$, pour les deux cas $\gamma \gg 1$ et $\gamma \ll 1$.
4. Discuter la stabilité des différentes solutions.

Discussion détaillée

1. Déterminer dans le plan (γ, δ), une représentation paramétrique de la ligne de séparation entre la phase magnétique et la phase non-magnétique.
2. Donner l'équation approchée de cette courbe pour $\gamma \rightarrow 0$ et pour $\gamma \rightarrow \frac{1}{\pi}$.
3. En déduire l'allure de la ligne de transition.
4. Exprimer la relation critique à l'aide de U et de la densité $\rho_{\mathrm{i},s}(\varepsilon_{\mathrm{F}})$.
5. δ étant fixé, tracer les variations des $\langle n_s \rangle$ dans la phase magnétique en fonction de γ.
6. Trouver les premiers termes du développement de la magnétisation pour γ petit, et son comportement au voisinage de la transition.
7. Pour $\delta = 1/4,\ 1/2$, tracer aussi précisément que possible la variation des $\langle n_s \rangle$ dans la phase non-magnétique en fonction de γ.

30.7.9 Tour d'horizon des propriétés magnétiques des solides

Il s'agit de rassembler et d'exploiter les idées fondamentales expliquant certaines propriétés magnétiques d'un solide, dont les objets constitutifs sont des atomes ou des ions à couches incomplètes. Tous les aspects diamagnétiques sont ignorés dans la suite.

La première partie est une analyse semi-quantitative montrant qu'une théorie à une particule ne peut décrire qu'un seul type de magnétisme (le paramagnétisme, sous une forme ou une autre), en contradiction avec l'expérience : la théorie des bandes est foncièrement inapte à expliquer le ferromagnétisme et sa caractéristique principale, l'existence d'une aimantation spontanée en champ magnétique nul au-dessous d'une certaine température T_c.

Dans la seconde partie, on rappelle comment la répulsion de Coulomb entre les électrons produit une forte dépendance de l'énergie d'un atome (ou d'un ion) par rapport au moment magnétique d'origine électronique. Puis, on introduit un Hamiltonien effectif reproduisant cette dépendance pour $N = 2$ électrons, qui est généralisé au cas d'un solide contenant $N \gg 1$ atomes ou ions (modèle de Heisenberg).

Dans la troisième partie, on traite le modèle de Heisenberg en champ moyen, ce qui permet de mettre en évidence simplement l'existence d'une température T_c au-dessous de laquelle il existe une magnétisation non nulle en champ nul (aimantation spontanée).

Propriétés magnétiques prévues par la théorie des bandes

Métaux Pour un métal (un alcalin, par exemple), on considère exclusivement la plus haute bande partiellement occupée, d'énergie $\varepsilon(\vec{k}\,)$, le niveau de Fermi ε_F se situant vers le milieu de la bande. m et e désignent la masse et la charge de l'électron, $\vec{S}$ représente son spin ($S = 1/2$), $\hbar M_S$ la valeur propre de S_z.

1. A la limite d'un échantillon macroscopique, toutes les sommes sur $\vec{k}$ peuvent être remplacées par des intégrales. Pour un intégrand ne dépendant que de l'énergie, on peut ainsi écrire :

$$\sum_{\vec{k}} \bullet = \int \mathrm{d}\varepsilon\, \mathcal{D}(\varepsilon) \bullet \;; \tag{30.286}$$

cette relation définit $\mathcal{D}(\varepsilon)$, densité d'états de la bande $\varepsilon(\vec{k})$ *compte non tenu du spin* ; quelle est l'expression de $\mathcal{D}(\varepsilon)$, $\mathcal{D}^{(0)}(\varepsilon)$, si l'on s'en tient à l'approximation $\varepsilon(\vec{k}\,) \simeq \varepsilon^{(0)}(\vec{k}\,) = \hbar^2\vec{k}^{\,2}/(2m)$? Exprimer $\mathcal{D}^{(0)}(\varepsilon_\mathrm{F})$ à l'aide de k_F pour un gaz d'électrons libres.

2. En présence d'un champ magnétique $\vec{\mathcal{B}}$ dirigé le long de Oz, le couplage entre un électron et ce champ est $-\mu_z\mathcal{B}$, où le moment magnétique $\vec{\mu}$ est relié au spin par :

$$\vec{\mu} = g_\mathrm{e}\frac{e}{2m}\,\vec{S}\ . \tag{30.287}$$

Quelle est l'énergie $E_{M_S}(\vec{k}\,)$ d'un électron ? Exprimer celle-ci notamment en fonction de M_S et du magnéton de Bohr $\mu_{\mathrm{B}} = e\hbar/(2m) \simeq -6\times 10^{-5}$ eV/T, en prenant $g_{\mathrm{e}} = 2$.

3. Les électrons sont répartis conformément à la loi de Fermi - Dirac :

$$f(E) = \frac{1}{\mathrm{e}^{\beta(E-\mu)}+1} \qquad (\beta = (k_{\mathrm{B}}T)^{-1}) \ , \tag{30.288}$$

où μ désigne le potentiel chimique. Le nombre moyen d'électrons de spin M est donc :

$$n_{M_S} = \int \mathrm{d}^3k\, f(E_{M_S}) \ . \tag{30.289}$$

Écrire n_{M_S} sous la forme d'une intégrale portant sur l'énergie.

4. Le potentiel chimique μ dépend du champ magnétique. Lorsque[88] $|\mu_{\mathrm{B}}|\mathcal{B} \ll \varepsilon_{\mathrm{F}}$, cette variation est du second ordre :

$$\mu(T,\,\mathcal{B}) = \mu(T,\,0)\left[1+\mathcal{O}\Big(\frac{\mu_{\mathrm{B}}\mathcal{B}}{\mu(T,\,0)}\Big)^2\right] \ ; \tag{30.290}$$

cette faible dépendance de μ est négligée dans la suite.

Soit $\mathcal{M}$ la magnétisation du métal (moment magnétique moyen de l'ensemble des électrons). Exprimer $\mathcal{M}$ en fonction des n_{M_S} et de μ_{B}, puis sous la forme d'une intégrale d'énergie contenant deux fonctions f décalées.

5. En utilisant le fait que $|\mu_{\mathrm{B}}|\mathcal{B} \ll \varepsilon_{\mathrm{F}}$, effectuer un développement limité pour en déduire l'expression de $\mathcal{M}$ à température nulle. Montrer que la susceptibilité magnétique est ici donnée par[89] :

$$\chi_{\mathrm{Pauli}} = 2\mu_{\mathrm{B}}^2\mathcal{D}(\varepsilon_{\mathrm{F}}) \ . \tag{30.292}$$

6. À température finie et dans des conditions usuelles, on a toujours $k_{\mathrm{B}}T,\ |\mu_{\mathrm{B}}|\mathcal{B} \ll \varepsilon_{\mathrm{F}}$. À quelle condition sur $\varepsilon(\vec{k}\,)$ la variation de χ_{Pauli} en température reste-t-elle très faible ?

Isolants Pour un isolant formé d'ions à couches incomplètes, on schématise la situation comme suit. D'une part la bande (pleine) la plus haute en énergie (peu dispersée), d'autre part des bandes *internes* sans dispersion, provenant d'états atomiques pratiquement inaltérés lors de la formation du solide.

[88] cas usuel !

[89] Comme définition générale de la susceptibilité magnétique, on adoptera :

$$\chi = \left(\frac{\partial\mathcal{M}}{\partial\mathcal{B}}\right)_{\mathcal{B}=0,\,T} \ . \tag{30.291}$$

Le résultat (30.292) décrit le paramagnétisme dit de Pauli.

1. L'observation fréquente d'un fort paramagnétisme s'explique par le fait que les ions, ayant une couche incomplète, peuvent posséder un moment magnétique moyen non-nul. Le couplage Zeeman entre l'ion et le champ est de la forme :

$$-\frac{e}{2m}(L_z + g_{\rm e}\, S_z)\mathcal{B} \ . \tag{30.293}$$

   On sait que, $\vec{J}$ étant le moment cinétique total, il existe (dans un sous-espace $\mathcal{E}_J$, J donné) une relation de proportionnalité entre les divers éléments de matrice, laquelle s'écrit précisément :

$$\langle L,\, S,\, J,\, M_J|L_z{+}g_{\rm e}\, S_z|L,\, S,\, J,\, M'_J\rangle = g\,\langle L,\, S,\, J,\, M_J|J_z|L,\, S,\, J,\, M'_J\rangle \ , \tag{30.294}$$

   où g est un nombre (facteur de Landé)[90]. Quelle est en présence du champ magnétique l'énergie $E_{M_J}^{(\rm ion)}$ d'un ion dans l'état $|L,\, S,\, J,\, M_J\rangle$?

2. En considérant exclusivement les niveaux d'un multiplet J, quelle est la probabilité P_{M_J} d'occupation de l'état $|L,\, S,\, J,\, M_J\rangle$ à la température T ?

3. En déduire l'expression formelle de la magnétisation totale moyenne $\mathcal{M}$ pour un échantillon contenant N ions. Montrer qu'elle est donnée par :

$$\mathcal{M} = N\, g\mu_{\rm B} J\, B_J(X) \ , \tag{30.296}$$

   où B_J est la fonction de Brillouin (voir Tome 1, problème 11.3.1) :

$$B_J(X) = \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}\,X\right) - \frac{1}{2J}\coth\left(\frac{1}{2J}\,X\right) \tag{30.297}$$

   et où $X = gJ\mu_{\rm B}\beta\mathcal{B}$.

4. Sachant que $\coth x \simeq (1/x) + (x/3)$ quand $x \ll 1$, en déduire la suceptibilité magnétique[91] $\chi_{\rm vV}$ due aux couches incomplètes.

5. Illustrer par un schéma la variation de $\chi_{\rm vV}$ en température (loi de Curie).

6. En assimilant $\mathcal{D}$ et $\varepsilon_{\rm F}$ à leurs expressions pour N électrons libres dans le volume V, comparer, à l'ambiante, les susceptibilités $\chi_{\rm Pauli}$ et $\chi_{\rm vV}$.

Hamiltonien effectif

2 électrons Soit un système à deux électrons de spins $\vec{S}_1$ et $\vec{S}_2$, soumis chacun à un certain champ de forces purement électrostatique[92] et en interaction mutuelle par l'interaction de Coulomb $V(r_{12})$. Le Hamiltonien est de la forme :

$$H = H^{(1)}(1) + H^{(1)}(2) + V(r_{12}) \equiv H_0 + V \ , \tag{30.298}$$

[90] L'expression de g est la suivante :

$$g = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \qquad \text{avec } g_{\rm e} = 2 \ . \tag{30.295}$$

[91] Cette forme de paramagnétisme est appelée paramagnétisme de van Vleck.

[92] On laisse de côté tout ce qui a trait à l'interaction spin-orbite.

où H_0 est relatif au modèle à particules indépendantes.

1. On désigne par $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, … les états propres de spin communs à S_{1z} et S_{2z}. Former les états propres de $(\vec{S}^2,\, S_z)$, $|S,\, M_S\rangle$, où $\vec{S}$ est le spin total $\vec{S} = \vec{S}_1 + \vec{S}_2$.
2. Soit $\{\phi_n\}$ un jeu d'orbitales propres de $H^{(1)}$. Quelles sont les fonctions d'onde d'espace acceptables construites sur deux orbitales *distinctes* ϕ_n et $\phi_{n'}$?
3. Deux telles orbitales étant fixées, on traite V par perturbation ; trouver la correction d'énergie au plus bas ordre pour les deux états possibles du spin total, en fonction des deux éléments de matrice :
$$\mathcal{J} = \langle \phi_n\phi_{n'}|V|\phi_n\phi_{n'}\rangle \ , \qquad \mathcal{K} = \langle \phi_n\phi_{n'}|V|\phi_{n'}\phi_n\rangle \ . \tag{30.299}$$
Dans quelle gamme d'énergie se situe la valeur numérique de ces intégrales ?
4. Expliquer physiquement pourquoi l'énergie (perturbée) de l'état triplet est *inférieure* à celle de l'état singulet. Exprimer la différence $E_\mathrm{S} - E_\mathrm{T}$ en fonction de $\mathcal{K}$.
5. Soit le Hamiltonien effectif :
$$H_\mathrm{eff} = -\hbar^{-2} J\, \vec{S}_1.\vec{S}_2 \ . \tag{30.300}$$
Sur quelle base H_eff est-il diagonal ? Quelles sont ses valeurs propres ?
6. Relier J défini dans l'équation (30.300) à $\mathcal{K}$ (équation (30.299)) de sorte que, à une constante additive près, H_eff et $H_0 + V$ aient le même spectre.

N électrons sur N sites d'un réseau On considère maintenant un réseau[93] de Bravais dont les nœuds sont repérés par $\vec{R}$; en chaque nœud se trouve un atome de spin $\vec{S}_{\vec{R}}$ ($S = 1/2$). On admet que ce système – au moins pour les excitations (magnétiques) de basse énergie – peut être représenté par un Hamiltonien effectif généralisant le cas $N = 2$, appelé *Hamiltonien de Heisenberg.* En champ magnétique nul, on adopte ainsi dorénavant :
$$H_\mathrm{Heis} = -\frac{1}{2\hbar^2}\sum_{\vec{R}}\ \sum_{\vec{R}'\neq\vec{R}} J(\vec{R}-\vec{R}')\, \vec{S}_{\vec{R}}.\vec{S}_{\vec{R}'} \ , \tag{30.301}$$
où $J(\vec{R}) > 0\ \forall\, \vec{R}$ (version *ferromagnétique* du modèle de Heisenberg).

1. L'état fondamental est l'état où les spins sont tous parallèles entre eux (chaque spin est dans l'état $|\uparrow\rangle$). Donner l'expression de l'énergie fondamentale E_0.
2. Décrire les *quanta* d'excitation élémentaires du réseau de Heisenberg, considéré comme une juxtaposition de ses N sites.
3. En pensant à l'approximation des liaisons fortes, contruire N états propres de H_Heis construits sur ces excitations de basse énergie.

[93]Comme toujours, on suppose que le système physique contient N sites dans un volume V, ($N \gg 1$).

Traitement en champ moyen du modèle de Heisenberg

Les excitations qui viennent d'être mises en évidence permettent de comprendre certaines propriétés à basse température. L'étude à température quelconque exige la connaissance de *toutes* les excitations élémentaires et est de ce fait hors de portée.

On adopte maintenant un point de vue complètement différent, en traitant le problème par une approximation de champ moyen. En se focalisant sur *un* spin donné (celui qui est en $\vec{R}$), on pose :

$$H_{\text{Heis}} = \sum_{\vec{R}} H_{\vec{R}} \ , \qquad H_{\vec{R}} = -\hbar^{-2}\vec{S}_{\vec{R}}.\Big(\frac{1}{2}\sum_{\vec{R}'} J(\vec{R}-\vec{R}')\,\vec{S}_{\vec{R}'}\Big) \ . \tag{30.302}$$

L'approximation de champ moyen consiste à remplacer dans (30.302) $\vec{S}_{\vec{R}'}$ par sa valeur moyenne (thermique) $\langle\vec{S}_{\vec{R}'}\rangle$. Dès lors, le spin singularisé est en simple interaction avec un champ effectif $\vec{\mathcal{B}}_{\text{mf}}$:

$$\vec{\mathcal{B}}_{\text{mf}} = \frac{1}{2g_{\text{e}}\hbar\mu_{\text{B}}} \sum_{\vec{R}'\neq\vec{R}} J(\vec{R}-\vec{R}')\,\langle\vec{S}_{\vec{R}'}\rangle \ , \tag{30.303}$$

et l'on se retrouve formellement dans la situation rencontrée dans la partie traitant des isolants, p. 1445, $\vec{\mathcal{B}}_{\text{mf}}$ jouant ici le rôle que joue $\vec{\mathcal{B}}$ là-bas. Dans les notations présentes, la magnétisation est :

$$\mathcal{M} = N\,g_{\text{e}}\frac{\mu_{\text{B}}}{\hbar}\,\langle\vec{S}_{\vec{R}}\rangle \ , \tag{30.304}$$

notant que, pour un échantillon homogène, la moyenne thermique $\langle\vec{S}_{\vec{R}}\rangle$ est indépendante de $\vec{R}$.

1. Écrire la relation satisfaite par $\mathcal{M}$ en utilisant la fonction de Brillouin $B_{J=1/2}$.
2. Discuter graphiquement la(es) solution(s) de l'équation précédente. Montrer notamment qu'il existe une certaine température notée T_{c} en-dessous de laquelle la magnétisation en champ nul est non-nulle. Démontrer que T_{c} est donnée par :

$$T_{\text{c}} = \frac{1}{4k_{\text{B}}}\sum_{\vec{R}} J(\vec{R}) \ . \tag{30.305}$$

3. Préciser les conditions physiques dans lesquelles l'approximation de champ moyen est *a priori* raisonnable.

30.7.10 Magnons et états liés de magnons dans le modèle de Heisenberg

Le modèle de Heisenberg vise à décrire certaines propriétés magnétiques des solides. Ce modèle généralise pour un ensemble d'atomes l'idée simple suivant laquelle les interactions

coulombiennes entre deux électrons peuvent être traduites par un Hamiltonien effectif du genre $\vec{S}_1.\vec{S}_2$ où les $\vec{S}_i$ désignent les moments cinétiques de spin des électrons.

Dans toute la suite, on considère un réseau unidimensionnel de pas a ; le Hamiltonien de Heisenberg pour une chaîne parallèle à Oy de N spins quantiques $\vec{S}_n = \frac{\hbar}{2}\vec{\sigma}_n$ s'écrit comme suit :

$$H_{\text{Heis}} = -\frac{J}{4}\sum_{n=1}^{N}\vec{\sigma}_n.\vec{\sigma}_{n+1} \qquad (N \gg 1) \tag{30.306}$$

Les opérateurs de deux spins distincts commutent entre eux. J est une quantité positive (cas ferromagnétique). En présence d'un champ magnétique $\vec{\mathcal{B}}$ parallèle à l'axe Oz, le Hamiltonien total H doit être complété par le couplage Zeeman :

$$H = H_{\text{Heis}} + H_{\text{Zeeman}} \;, \qquad H_{\text{Zeeman}} = -\frac{1}{2}\hbar\omega_{\text{c}}\sum_{n=1}^{N}\sigma_{n\,z} \;. \tag{30.307}$$

Le cas échéant, on invoquera des conditions cycliques de Born - von Kármán, en identifiant $N+1$ et 1. On rappelle l'expression des matrices de Pauli sur la base $|+\rangle$, $|-\rangle$ des deux états propres de $(\vec{S}^{\,2},\, S_z)$ d'un spin $S = \frac{1}{2}$:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \;, \quad \sigma_y = \begin{bmatrix} 0 & -\text{i} \\ \text{i} & 0 \end{bmatrix} \;, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \;, \qquad \sigma_{\pm} = \frac{1}{2}(\sigma_x \pm \text{i}\sigma_y) \;, \tag{30.308}$$

et les relations caractéristiques :

$$[\sigma_x,\, \sigma_y] = 2\text{i}\sigma_z, \ldots \;, \qquad \{\sigma_u,\, \sigma_v\} = 2\delta_{uv}\mathbf{1} \;, \quad \sigma_{\pm}|\pm\rangle = 0 \;, \quad \sigma_{\pm}|\mp\rangle = |\pm\rangle \;. \tag{30.309}$$

On pourra utiliser la relation :

$$\vec{\sigma}_n.\vec{\sigma}_{n'} = \sigma_{nz}\sigma_{n'z} + 2(\sigma_{n+}\sigma_{n'-} + \sigma_{n-}\sigma_{n'+}) \;. \tag{30.310}$$

États à une excitation au plus

1. Décrire l'état fondamental, noté $|\text{vac}\rangle$ dans toute la suite. Quelle est son énergie E_0 en présence du champ ? Dans toute la suite, les énergies sont comptées à partir de E_0, ce qui revient à poser $E_0 = 0$.

2. Soit $|p\rangle$ l'état obtenu en basculant le spin $\vec{S}_p$ de l'état fondamental :

$$|p\rangle = \sigma_{-p}|\text{vac}\rangle \;. \tag{30.311}$$

   L'ensemble des $\{|p\rangle\}_p$ engendre un sous-espace noté $\mathcal{E}_1$.

   (a) Calculer $\sigma_{nz}\sigma_{n+1z}|p\rangle$, puis $\sigma_{n-}\sigma_{n+1+}|p\rangle$.

   (b) En déduire $H|p\rangle$.

3. Soit $\mathcal{T}$ l'opérateur agissant dans l'espace des états et associé à une translation élémentaire du réseau ; on a en particulier :

$$\mathcal{T}|p\rangle = |p+1\rangle \ . \tag{30.312}$$

Soit $|k\rangle$ un vecteur propre de $\mathcal{T}$ appartenant à $\mathcal{E}_1$, $\mathcal{T}|k\rangle = \mathrm{e}^{-\mathrm{i}ka}|k\rangle$.

   (a) Rappeler pourquoi k doit être réel.
   (b) En raisonnant avec $\mathcal{T}^N$, trouver les valeurs possibles de k.
   (c) On pose $|k\rangle = \sum_p a_{kp}|p\rangle$. Trouver la relation entre a_{kp} et a_{kp-1}.
   (d) D'une phrase, expliquer pourquoi les états $|k\rangle$ sont également états propres de H. Donner l'expression normalisée de ces derniers.

4. Déduire de ce qui précède les valeurs propres $\varepsilon(k)$ de H_{Heis} dans le sous-espace $\mathcal{E}_1$.

5. Chaque état $|k\rangle$ est une excitation au-dessus du fondamental appelée *magnon.*

   (a) Tracer la relation de dispersion de ces magnons dans l'intervalle $[-\pi, +\pi]$, qui joue le rôle de première zone de Brillouin .
   (b) Calculer les valeurs moyennes dans un état $|k\rangle$ des produits $\sigma_{nz}\sigma_{n'z}$, $\sigma_{nx}\sigma_{n'x}$ et $\sigma_{nx}\sigma_{n'y}$.
   (c) En raisonnant par référence à un spin donné, $\vec{\sigma}_{n=0}$ par exemple, montrer qu'un état de magnon peut être schématisé comme une chaîne de spins précessant dans l'espace autour de Oz quand on parcourt la chaîne.
   (d) Préciser la structure magnétique quand k est en centre de zone ($k=0$) ou en bord de zone ($k = \pi/a$).

6. Établir (30.217).

États liés de deux magnons

Dans cette partie, on considère les états appartenant au sous-espace $\mathcal{E}_2$ engendré par tous les vecteurs $|pp'\rangle$ formés à partir du fondamental en basculant les deux spins situés en p et p' :

$$|pp'\rangle = \sigma_{-p}\sigma_{-p'}|\text{vac}\rangle \ . \tag{30.313}$$

1. Combien y a-t-il d'états $|pp'\rangle$?

2. Expliquer (sans calculs) que $\mathcal{E}_2$ est stable par H.

3. Soit les états à deux magnons indépendants, notés $|kk'\rangle$, définis comme le simple produit tensoriel :

$$|kk'\rangle \equiv |k\rangle \otimes |k'\rangle = C \sum_{p,p'=1}^{N} \mathrm{e}^{\mathrm{i}(kp+k'p')a}\,|pp'\rangle \ , \tag{30.314}$$

où C est une constante de normalisation.

(a) Combien y a-t-il d'états $|kk'\rangle$ distincts ?
(b) En déduire que ces états ne peuvent être propres de H.

4. Il s'agit maintenant de trouver l'action de H sur un état du type $|pp'\rangle$.

(a) De combien varie l'énergie lorsque l'on retourne les deux spins situés en p et p' (distinguer les cas $|p-p'| \neq 1$ et $|p-p'| = 1$) ? En déduire que les éléments diagonaux de la matrice de H sur la base $|pp'\rangle$ sont égaux à :

$$2\hbar\omega_c + J[2 - (\delta_{p'p+1} + \delta_{p'p-1})] \ . \qquad (30.315)$$

(b) Montrer que $\sigma_{+n+1}|p\,p'\rangle = \delta_{n+1p}|p'\rangle + \delta_{n+1p'}|p\rangle$.
(c) En notant que $\sigma_{-\,n}|p\rangle = (1-\delta_{np})|np\rangle$, déduire de 4b que :

$$\sum_n \sigma_{-\,n}\sigma_{+n+1}|pp'\rangle = |p-1p'\rangle + |pp'-1\rangle \ . \qquad (30.316)$$

5. L'analyse précédente montre que deux magnons libres $|k\rangle$ et $|k'\rangle$ sont en interaction ; l'existence d'états liés, établie ci-dessous, montre que cette interaction est attractive. On peut donc visualiser la situation en imaginant deux magnons d'impulsions k_1 et k_2 subissant des collisions.

(a) Combien vaut $\mathcal{T}|pp'\rangle$?
(b) En déduire que $\mathcal{T}|kk'\rangle = \mathrm{e}^{-\mathrm{i}(k+k')a}|kk'\rangle$.
(c) En raisonnant avec le commutateur $[H_{\mathrm{Heis}}, \mathcal{T}]$, montrer que H_{Heis} ne couple que les états $|k_1\,k_2\rangle$ et $|k_1'\,k_2'\rangle$ tels que $k_1+k_2 = k_1'+k_2'$. Interpréter ce résultat.

6. Soit $|\Psi\rangle$ un état propre de H appartenant à $\mathcal{E}_2$, $H|\Psi\rangle = E|\Psi\rangle$, développé sur la base $|pp'\rangle$: $|\Psi\rangle = \sum_{p,p'} c_{pp'}\,|pp'\rangle$. Compte tenu des résultats acquis en 5, on peut poser :

$$c_{pp'} = \mathrm{e}^{\mathrm{i}K\frac{p+p'}{2}a}\,\psi_K(p-p') \ , \qquad (30.317)$$

$K = k_1 + k_2$ est l'impulsion totale de deux magnons. Expliquer pourquoi $\psi_K(p-p')$ est une fonction paire. Combien vaut $\psi_K(0)$?

7. Les résultats établis en 4. montrent que les coefficients c_{p+rp} sont donnés par le système :

$$[\varepsilon + 2(\delta_{r\,+1} + \delta_{r\,-1})]\,c_{p+rp} + c_{p+r+1p} + c_{p+r-1p} + c_{p+rp+1} + c_{p+rp-1} = 0 \ , \qquad (30.318)$$

avec $r \neq 0$ et où $\varepsilon = 2\frac{E-2\hbar\omega_c-2J}{J}$. Dans la suite, on suppose $-\pi \leq K \leq +\pi$.

(a) En déduire, pour $r \geq 1$, la relation satisfaite par la fonction $\psi_K(r)$ définie en (30.317).
(b) On pose $\psi_K(r) = \mathrm{e}^{-ru}$ $(r \geq 1,\ u > 0)$. À l'aide de la relation écrite en 7a, montrer que[94] :

$$\varepsilon + 4\cos\frac{Ka}{2}\cosh u = 0 \ , \qquad -\frac{\varepsilon+2}{2\cos\frac{Ka}{2}} = \mathrm{e}^{-u} \ . \qquad (30.319)$$

[94] Considérer successivement les deux cas $r \geq 2$ et $r = 1$.

(c) En déduire ε, puis l'énergie propre correspondante E_b.

8. Quelle est l'énergie $E(K, q)$ d'une paire de magnons $|k = \frac{K}{2} + q,\, k' = \frac{K}{2} - q\rangle$ sans interaction ?

9. Montrer que $E_b \leq \min_q E(K, q)$ et en déduire l'énergie de liaison E_{liaison} de la paire de magnons liés. Tracer E_{liaison} en fonction de K.

10. La fonction d'onde de l'état lié, $\psi_K(r)$, est une exponentielle caractérisée par une longueur ξ. Montrer que :
$$\xi = -\frac{a}{\ln\cos\frac{Ka}{2}} \ , \tag{30.320}$$
Corréler la variation de ξ à celle de l'énergie de liaison. Exprimer à l'aide de $\psi_K(r)$ la valeur moyenne dans l'état lié de la distance entre les deux spins retournés ; donner l'allure de sa variation en fonction de K.

30.7.11 Excitations d'un réseau de spins

Il s'agit d'examiner les propriétés magnétiques d'un condensat de spins et notamment d'expliquer l'écart aux lois classiques. Lorsque les spins sont sans interaction, le système se comporte comme un paramagnétique parfait, dont la susceptibilité obéit à une loi de Curie ordinaire, simplement revisitée pour tenir compte de la quantification des moments magnétiques (cinétiques).

Expérimentalement, on trouve souvent que, à basse température, la magnétisation en champ nul à la température T suit en fait une loi du genre :
$$\mathcal{M}(T) = [1 - \mathrm{C}^{\text{ste}}\, T^{3/2}]\, \mathcal{M}(0) \ . \tag{30.321}$$

Cette variation peut s'expliquer à l'aide d'un modèle prenant en compte des interactions entre les spins. Dans la suite, on s'intéresse à l'état fondamental et aux excitations de basse énergie (magnons) dans le cadre du modèle de Heisenberg ; l'origine et les fondements de celui-ci sont établis dans la première partie. Dans la deuxième, les états pertinents de ce modèle sont construits en détail. Enfin, on démontre que la magnétisation obtenue dans ce cadre obéit bien à une loi du genre de celle donnée en (30.321).

Deux spins

Soit un système de deux spins 1/2 qui commutent, $\vec{S}_1$ et $\vec{S}_2$. Les états propres de S_{1z}, S_{2z}, $\vec{S}_1^{\,2}$, $\vec{S}_2^{\,2}$ sont conventionnellement notés $|\varepsilon_1\, \varepsilon_2\rangle$, où $\varepsilon = \pm 1$ donne le signe de la valeur propre de la projection le long de Oz.

1. Quelle est l'interaction la plus simple invariante par rotation des spins (justifier la réponse) ?

2. Calculer tous les vecteurs du genre $\vec{S}_1.\vec{S}_2\,|\varepsilon_1\,\varepsilon_2\rangle$.

3. En déduire l'égalité :

$$\hbar^{-2}\vec{S}_1.\vec{S}_2|\varepsilon_1\,\varepsilon_2\rangle = -\frac{1}{4}|\varepsilon_1\,\varepsilon_2\rangle + \frac{1}{2}|\varepsilon_2\,\varepsilon_1\rangle \ . \tag{30.322}$$

4. Que fait l'opérateur $2\hbar^{-2}\vec{S}_1.\vec{S}_2 + \frac{1}{2}\mathbf{1}$? Quel nom proposez-vous de lui donner ?

N spins sur réseau

On considère maintenant un ensemble de spins 1/2, dont chacun est situé en un nœud d'un réseau de Bravais cubique simple, $\mathcal{B}$. Le spin localisé en $\vec{R}$ est noté $\vec{S}_{\vec{R}}$; d'un site à l'autre, tous les opérateurs commutent entre eux :

$$[\vec{S}_{\vec{R}},\, \vec{S}_{\vec{R}'}] = 0 \quad \forall\, \vec{R} \neq \vec{R}' \ . \tag{30.323}$$

Comme d'habitude, on suppose le réseau macroscopique mais fini ; N désigne le nombre de spins, z est le nombre de premiers voisins d'un nœud donné. Le cas échéant, on invoquera des conditions aux limites périodiques.

En présence d'interactions, et au vu des résultats précédents, on écrit le Hamiltonien-modèle le plus simple sous la forme d'une somme de termes d'échange (modèle de Heisenberg) :

$$H = -\hbar^{-2}J \sum_{\langle \vec{R},\, \vec{R}'\rangle} \vec{S}_{\vec{R}}.\vec{S}_{\vec{R}'} \qquad (J > 0) \ . \tag{30.324}$$

Le symbole $\langle \vec{R},\, \vec{R}'\rangle$ signifie que la sommation est faite sur toutes les paires ordonnées de deux spins situés aux deux nœuds $\vec{R}$ et $\vec{R}'$ qui sont premiers voisins l'un de l'autre.

État fondamental

1. Quelle est l'énergie d'interaction de deux spins premiers voisins de même orientation ?

2. Expliquer pourquoi, dans l'état fondamental (noté $|0\rangle$), tous les spins sont parallèles.

3. Quel est, en fonction de N et de z, le nombre de paires ?

4. Quelle est l'énergie E_0 de l'état fondamental.

Excitations de basse énergie Pour construire les premiers états excités, on considère les N états déduits de $|0\rangle$ en retournant un seul spin. Afin de fixer les idées, on prend pour $|0\rangle$ l'état où tous les ε sont égaux à $+1$; dès lors, il existe N états notés $|\vec{R}\rangle$ où tous les spins pointent vers le bas, à l'exception de celui sur le site $\vec{R}$, qui pointe vers le haut.

1. D'une phrase, expliquer pourquoi la question de l'indiscernabilité ne se pose pas dans le présent contexte.
2. En raisonnant physiquement, trouver la différence $\langle\vec{R}|H|\vec{R}\rangle - E_0$.
3. Quel est le nombre de paires n'impliquant pas le spin basculé ?
4. Que vaut $\vec{S}_{\vec{R}'}.\vec{S}_{\vec{R}''}|\vec{R}\rangle$ quand $\vec{R}' \neq \vec{R}$ et $\vec{R}'' \neq \vec{R}$?
5. Que vaut $\vec{S}_{\vec{R}'}.\vec{S}_{\vec{R}''}|\vec{R}\rangle$ quand $\vec{R}' = \vec{R}$ ou $\vec{R}'' = \vec{R}$?
6. En déduire l'égalité suivante :
$$H|\vec{R}\rangle = -\big(\frac{N}{2} - 1\big)\,\frac{zJ}{4}\,|\vec{R}\rangle - \frac{J}{2}\sum_{\vec{R}'\in\mathcal{V}(\vec{R})}\Big(|\vec{R}'\rangle - \frac{1}{2}|\vec{R}\rangle\Big)\ , \qquad (30.325)$$
où $\mathcal{V}(\vec{R})$ désigne l'ensemble des spins voisins du spin basculé.
7. Simplifier au maximum le second membre de (30.325), notamment en y faisant apparaître E_0.
8. Soit l'état $|\vec{k}\rangle$:
$$|\vec{k}\rangle = N^{-1/2}\sum_{\vec{R}\in\mathcal{B}} \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}}|\vec{R}\rangle\ . \qquad (30.326)$$
Cet état satisfait-il le théorème de Bloch ?
9. En argumentant à partir de la réponse à la question précédente, expliquer pourquoi l'égalité suivante est vraie :
$$H|\vec{k}\rangle = E(\vec{k}\,)\,|\vec{k}\rangle\ . \qquad (30.327)$$
L'état $|\vec{k}\rangle$ est une excitation élémentaire de type onde de spin (*magnon*).
10. Établir l'expression suivante de l'énergie $E(\vec{k}\,)$:
$$E(\vec{k}\,) = E_0 + \frac{J}{2}\sum_{\vec{R}\in\mathcal{V}(\vec{0})}(1 - \mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}})\ . \qquad (30.328)$$
En déduire que l'énergie d'un magnon est :
$$E(\vec{k}\,) = E_0 + J\sum_{\vec{R}\in\mathcal{V}(\vec{0})}\sin^2\frac{\vec{k}.\vec{R}}{2} \equiv E_0 + J\Delta(\vec{k}\,)\ . \qquad (30.329)$$

Variation de la magnétisation à basse température

La magnétisation $\mathcal{M}$ est :

$$\mathcal{M} = g_{\rm e}\mu_{\rm B}\hbar^{-1}\langle S_z\rangle \ , \qquad \mu_{\rm B} = \frac{e\hbar}{2m} \ , \tag{30.330}$$

où S_z est la composante suivant Oz du spin total $\vec{S} = \sum_{\vec{R}\in\mathcal{B}} \vec{S}_{\vec{R}}$.

1. Montrer que $|\vec{k}\rangle$ est vecteur propre de S_z et donner la valeur propre correspondante.
2. Quelle est la conséquence de l'excitation thermique des magnons sur la valeur moyenne de l'aimantation ?
3. Pour pouvoir calculer les moyennes thermiques, il convient au préalable d'élucider la nature d'un magnon : boson ou fermion ? Pour cela, on considère les états à deux magnons qui sont certaines combinaisons linéaires des états obtenus en basculant *deux* spins à partir du fondamental ; un tel état est noté $|\vec{R}, \vec{R}'\rangle$.
   (a) Expliquer pourquoi $|\vec{R}, \vec{R}'\rangle = |\vec{R}', \vec{R}\rangle$, et conclure.
   (b) Peut-on créer une double excitation sur le même site $\vec{R}$?
   (c) Conclure en complètant la phrase :

   $$\text{Les magnons sont des } \ldots\ldots \text{ de cœur } \ldots \ . \tag{30.331}$$

   Dans la suite, on admettra que les magnons sont suffisamment dilués pour former un gaz parfait de particules sans interactions.
   (d) Quel est le spin d'un magnon ?
4. La magnétisation à la température T est donnée par :

   $$\mathcal{M}(T) = \mathcal{M}(0) + g_{\rm e}\mu_{\rm B}\sum_{\vec{k}} n_{\vec{k}}(T) \ , \tag{30.332}$$

   où $n_{\vec{k}}(T)$ est le nombre moyen de magnons à la température T :

   $$n_{\vec{k}}(T) = \frac{1}{\mathrm{e}^{\beta J\Delta(\vec{k})} - 1} \ . \tag{30.333}$$

   Commenter l'expression (30.332). Combien vaut $\mathcal{M}(0)$?
5. Montrer que $\mathcal{M}(T)$ peut se mettre sous la forme :

   $$\mathcal{M}(T) = \mathcal{M}(0)[1 - I(T)] \ , \tag{30.334}$$

   où $I(T)$ est une certaine intégrale à préciser.
6. On suppose maintenant que la température est suffisamment basse pour que seules les petites valeurs du module de $\vec{k}$ soient importantes. Montrer qu'alors l'intégrale $I(T)$ est de la forme :

   $$I(T) \simeq AT^{3/2} \ , \tag{30.335}$$

   où A est une constante pratiquement indépendante de la température. Conclusion ?

30.7.12 Boîte quantique sous champ magnétique

Depuis quelques années, on sait fabriquer des systèmes de taille nanométrique, appelées *boîtes quantiques* (*quantum dots*), qui permettent de confiner les électrons de conduction d'un solide à basse température. Une boîte quantique est réalisée en déposant un matériau A, jouant le rôle de puits, sur un substrat d'un autre matériau, B. La possibilité de contrôler les états d'énergie d'une boîte, notamment en jouant sur sa géométrie, permet des applications intéressantes en micro- et opto-électronique.

Dans la première partie, on analyse le mouvement bidimensionnel d'un électron dans une boîte constituant un potentiel harmonique. Dans la deuxième, on étudie les modifications apportées par l'application d'un champ magnétique. La troisième est une brève confrontation avec les résultats de l'expérience et propose une amélioration du modèle développé afin de rendre compte des observations.

L'existence de la structure de bande précise au sein du matériau A compliquant considérablement le formalisme, on admettra partout qu'un électron dans la boîte ne diffère d'un électron libre que par sa masse effective, notée μ.

Oscillateur harmonique à deux dimensions

On admet que l'énergie potentielle effective d'un électron confiné dans le plan xOy est :

$$V(x,\, y) = \frac{1}{2}\mu\omega^2(x^2 + y^2) \ , \tag{30.336}$$

de sorte que le Hamiltonien décrivant le mouvement bidimensionnel de l'électron est :

$$H_{\mathrm{h}} = \frac{p_x^2}{2\mu} + \frac{p_y^2}{2\mu} + V(x,\, y) \ \equiv\ H(x,\, p_x) + H(y,\, p_y) \ . \tag{30.337}$$

1. On définit les opérateurs d'annihilation a_u $(u = x,\, y)$ comme suit :

$$a_u = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{\mu\omega}{\hbar}}\,u + \frac{\mathrm{i}}{\sqrt{\hbar\mu\omega}}\,p_u\right) \ ; \tag{30.338}$$

les opérateurs de création $a_u^\dagger$ sont les hermitiques conjugués des a_u ; on a :

$$[a_u,\, a_v^\dagger] = \delta_{uv} \ , \qquad [a_u,\, a_v] = [a_u^\dagger,\, a_v^\dagger] = 0 \tag{30.339}$$

Exprimer $H(u,\, p_u)$ à l'aide de a_u et $a_u^\dagger$.

2. L'opérateur $N_u = a_u^\dagger a_u$ a pour valeurs propres les entiers naturels ; ses états propres sont notés $|n_u\rangle$:

$$N_u|n_u\rangle = n_u|n_u\rangle \qquad (n_u \in \mathbb{N}) \ . \tag{30.340}$$

Comment construit-on les états propres $|n_x,\, n_y\rangle$ de H_{h} ? Posant $n = n_x + n_y$, quelles sont les valeurs propres E_n correspondantes ? Quelle est la dégénérescence g_n de E_n ?

3. Soit L_z la composante du moment cinétique de l'électron sur l'axe Oz perpendiculaire au plan xOy. Exprimer L_z en fonction des a_u et $a_u^\dagger$.

4. On sait que :

$$a_u|n_u\rangle = \sqrt{n_u}\,|n_u - 1\rangle \ , \qquad a_u^\dagger|n_u\rangle = \sqrt{n_u + 1}\,|n_u + 1\rangle \ . \tag{30.341}$$

   Quel est le résultat de l'action de L_z sur un état $|n_x,\, n_y\rangle$? La composante L_z a-t-elle une valeur bien définie dans cet état ?

5. Par un argument de symétrie, montrer que $[H_\mathrm{h},\, L_z] = 0$.

6. Soit les opérateurs d'annihilation "gauche" et "droit" :

$$a_\mathrm{g} \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}(a_x + \mathrm{i}a_y) \ , \qquad a_\mathrm{d} \stackrel{\text{déf}}{=} \frac{1}{\sqrt{2}}(a_x - \mathrm{i}a_y) \tag{30.342}$$

   ainsi que leurs hermitiques conjugués $a_\mathrm{g}^\dagger$, $a_\mathrm{d}^\dagger$. Quelles sont les relations de commutation entre ces quatre opérateurs ?

7. Exprimer H_h et L_z en fonction de $N_\mathrm{g} = a_\mathrm{g}^\dagger a_\mathrm{g}$ et $N_\mathrm{d} = a_\mathrm{d}^\dagger a_\mathrm{d}$; combien vaut $[N_\mathrm{g},\, N_\mathrm{d}]$? En déduire que les vecteurs propres communs à N_g et N_d, également notés $|n_\mathrm{g},\, n_\mathrm{d}\rangle$, sont propres de H_h et L_z. n_g et n_d étant donnés, le vecteur $|n_\mathrm{g},\, n_\mathrm{d}\rangle$ s'exprime en combinaison linéaire de certains $|n_x,\, n_y\rangle$: lesquels ?

8. On note $E_{n=n_\mathrm{g}+n_\mathrm{d}}$ et $m\hbar$ les valeurs propres de H_h et L_z. Quelles sont les valeurs possibles de m dans un état d'énergie E_n ?

9. Pour un oscillateur à une dimension, la fonction d'onde fondamentale est la gaussienne :

$$\psi_0(x) = \left(\frac{\mu\omega}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-\frac{\mu\omega x^2}{2\hbar}} \ . \tag{30.343}$$

   Quelle est son extension spatiale Δx pour $\mu = 0{,}07\ m_\mathrm{e}$ (m_e est la masse de l'électron) et $\hbar\omega = 0{,}06$ eV ? À la température $T = 10$ K, quel est le rapport des populations entre les deux premiers états de l'oscillateur ? Quelle est la longueur d'onde d'absorption du rayonnement entre deux états consécutifs de l'oscillateur ?

Boîte quantique dans un champ magnétique

On applique maintenant un champ magnétique de module $\mathcal{B}$, dirigé le long de l'axe Oz ; le potentiel-vecteur $\vec{A}$ est choisi sous la forme :

$$\vec{A} = \frac{1}{2}\vec{\mathcal{B}} \times \vec{r} \ . \tag{30.344}$$

La fréquence cyclotron est définie comme $\omega_\mathrm{c} \stackrel{\text{déf}}{=} |e|\mathcal{B}/\mu$, et on pose $\tilde{\omega} \stackrel{\text{déf}}{=} \sqrt{\omega_\mathrm{c}^2/4 + \omega^2}$.

1. Dans le Hamiltonien (30.337), par quoi doit-on remplacer $\vec{p}$? On obtient ainsi un nouveau Hamiltonien noté $H_\mathcal{B}$.

2. Transformer l'expression de $H_{\mathcal{B}}$ pour y faire apparaître L_z.

3. En déduire que l'on peut toujours trouver des états propres communs à l'énergie totale et à la projection L_z.

4. Donner l'expression de l'énergie, $E_{n_\mathrm{g},\,n_\mathrm{d}}$. La dégénérescence en champ nul est-elle complètement levée en présence du champ ? Donner ses expressions approchées en champ faible ($\omega_\mathrm{c} \ll \omega$) et en champ fort ($\omega_\mathrm{c} \gg \omega$). Pour les trois premiers états, tracer la variation de leur énergie en fonction de $\mathcal{B}$. Pour quelle valeur de $\mathcal{B}$ les deux niveaux $E_{0,\,1}$ et $E_{2,\,0}$ se croisent-ils (en donner la valeur numérique) ?

5. À $T = 10\,\mathrm{K}$ et pour des champs magnétiques plus petits que $20\,\mathrm{T}$, de quel état partent les transitions en absorption ?

6. Trouver les expressions des fréquences $\nu_\pm$ ($\nu_- < \nu_+$) des deux premières transitions et tracer leur variation en fonction du champ magnétique.

7. On a jusqu'à présent supposé que le mouvement de l'électron est strictement bidimensionnel ; il s'agit maintenant de préciser les conditions justifiant cette hypothèse.

   (a) Pour un potentiel de confinement dans la direction Oz tel que $V_z(z) = 0$ si $0 < z < D$ et infini autrement, quels sont les niveaux d'énergie associés au mouvement suivant Oz ?

   (b) Quelle condition de type énergétique doit être satisfaite pour que seul le fondamental suivant Oz soit pertinent ? En déduire une inégalité entre D et Δx.

Comparaison avec l'expérience

Expérimentalement, on observe que les deux fréquences $\nu_\pm$ varient en fonction du champ magnétique ($\mathcal{B} \lesssim 10$ T) comme indiqué sur la figure 30.18 et que, par ailleurs, l'absorption dépend de la polarisation du champ électromagnétique. Il s'agit de rendre compte du comportement observé (à petit champ magnétique) en incorporant une anisotropie du potentiel dans le plan xOy. Pour simplifier, on se bornera à considérer le cas du champ nul ($\mathcal{B} = 0$).

1. On remplace désormais $V(x,\,y)$ par $\tilde{V}(x,\,y)$ défini comme :

$$\tilde{V}(x,\,y) \stackrel{\text{déf}}{=} \frac{1}{2}\mu\omega^2[(1+\varepsilon)x^2 + (1-\varepsilon)y^2] \;, \qquad (0 < \varepsilon \ll 1) \;. \tag{30.345}$$

   Dans ces conditions, le Hamiltonien du mouvement dans le plan est :

$$\tilde{H}_\mathrm{h} = H_\mathrm{h} + \varepsilon W \;. \tag{30.346}$$

   Quelle est l'expression de l'opérateur W ?

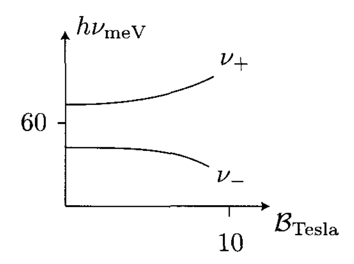


Figure 30.18: Variation observée des deux premières fréquences d'absorption en fonction du champ magnétique.

2. On traite l'anisotropie par perturbations au premier ordre des états $|n_x,\, n_y\rangle$ avec $n_x + n_y = 1$. Sans faire vraiment de calculs, indiquer l'allure de la matrice de perturbation dans le sous-espace dégénéré.

3. Calculer effectivement les éléments de matrice de la perturbation et en déduire les corrections d'énergie en champ nul résultant de l'anisotropie.

4. L'écart de fréquence mesuré en $\mathcal{B} = 0$ est voisin de 8 meV$/h$. En déduire la valeur numérique du taux d'anisotropie ε.

30.7.13 Oscillations de Bloch pour un atome

Il s'agit de mettre en évidence très directement le comportement apparemment paradoxal d'une particule soumise à un potentiel périodique.

On se souvient qu'un électron dans un métal suit une réflexion (dite de Bragg) lorsque son impulsion[95] atteint la valeur $\hbar\vec{K}/2$, où $\vec{K}$ est un vecteur du réseau réciproque $\tilde{\mathcal{B}}$. Pour les électrons dans un métal, l'observation directe est rendue impossible par le fait que le temps entre deux collisions est extrêmement court, de l'ordre de 10^{-14} s.

Le traitement ci-dessous s'inspire d'expériences faites avec des atomes ultra-froids [249], qui manifestent la même "anomalie" de comportement. Pour la simplicité, le mouvement de l'atome est supposé unidimensionnel ; sa coordonnée est notée x, le moment conjugué est p. Éventuellement, on bornera $\mathbb{R}$ au grand segment $[-L/2,\, +L/2]$, sur lequel les fonctions d'onde – sauf avis contraire – sont normalisées à l'unité.

Transposition à l'atome des résultats connus pour les électrons

Un atome (de masse M) est soumis à une onde stationnaire non-résonnante[96] de nombre d'onde k_0. En raison de sa polarisabilité, l'atome est soumis à une force proportionnelle

[95] prise au sens de *quasi-impulsion*.

[96] i.e. la fréquence de l'onde est loin de toute résonance atomique propre.

au *carré* du champ électrique et, au total, la force exercée par l'onde sur l'atome est proportionnelle à l'intensité de celle-ci. Le Hamiltonien de l'atome couplé semi-classiquement au champ est :

$$H_0 = \frac{p^2}{2M} + V_0 \sin^2(k_0 x) \qquad (V_0 > 0) \ . \tag{30.347}$$

1. Quelle est la période spatiale λ_0 du réseau optique ? Dans toute la suite, on pourra poser $b = 2k_0$ et, le cas échéant, on notera K tout vecteur d'onde de la forme nb avec $n \in \mathbb{Z}$.

2. Soit $\mathcal{T}(\lambda_0)$ l'opérateur de translation spatiale élémentaire, dont l'action sur toute fonction $f(x)$ est :

$$\mathcal{T}(\lambda_0) f(x) = f(x - \lambda_0) \equiv \mathrm{e}^{-\lambda_0 \frac{\mathrm{d}}{\mathrm{d}x}} f(x) \ . \tag{30.348}$$

   Exprimer $\mathcal{T}(\lambda_0)$ en fonction de l'impulsion p.

3. Le théorème de Bloch permet d'écrire toute fonction propre de H_0, $\psi_k(x)$, d'énergie $\varepsilon(k)$, sous la forme :

$$\psi_k(x) = \mathrm{e}^{\mathrm{i}kx} u_k(x) \ . \tag{30.349}$$

   En utilisant le fait que ψ_k est aussi fonction propre de $\mathcal{T}(\lambda_0)$ (quelle est la valeur propre correspondante ?), retrouver précisément la propriété fondamentale de la fonction u_k.

4. Deux fonctions ψ_k et ψ_{k+K} représentent-elles deux états différents ? Dans toute la suite, sauf mention contraire, on restreindra k à la première zone $[-k_0, +k_0]$.

5. Écrire l'équation aux valeurs propres satisfaites par la fonction u_k, en introduisant un Hamiltonien $H(x, p; k)$.

6. On désigne par :

$$\varepsilon^{(0)}(q) = \frac{\hbar^2 q^2}{2M} \tag{30.350}$$

   la fonction donnant la loi de dispersion d'un atome libre. Tracer la variation en fonction de k des valeurs propres de H_0 dans la limite $V_0 \to 0$.

7. En supposant V_0 petit, on traite le problème par perturbation, en ne considérant que les deux bandes non perturbées les plus basses en énergie, d'énergies $\varepsilon_n^{(0)}(k)$ définies comme suit :

$$\varepsilon_n^{(0)}(k) = \varepsilon^{(0)}(k - nb) \ , \quad n = 0, 1 \ , \quad k > 0 \ ; \qquad \varepsilon_n^{(0)}(-k) = \varepsilon_n^{(0)}(k) \ . \tag{30.351}$$

   Les fonctions correspondantes à l'ordre zéro sont notées $\psi_{n\,k}^{(0)}$; pour $k > 0$, on a :

$$\psi_{n\,k}^{(0)}(x) = \frac{1}{\sqrt{L}} \mathrm{e}^{\mathrm{i}(k-nb)x} \equiv \langle x | k - nb \rangle \ . \tag{30.352}$$

   Pour le traitement de perturbation, on envisage successivement les deux cas suivants :

(i) k est très loin de k_0. Calculer en fonction de l'amplitude V_0 la correction d'énergie au premier ordre, ΔE.

(ii) k est au voisinage de k_0. Trouver les éléments de matrice de la pertubation dans ce sous-espace et en déduire, par perturbation au premier ordre, les deux valeurs propres en présence du potentiel périodique. Représenter leur variation en fonction de k.

Introduction d'une force constante

On ajoute une force constante $F = Mg$; le Hamiltonien de l'atome devient :

$$H = H_0 - Mgx \qquad (g > 0) \ . \tag{30.353}$$

1. La solution de l'équation de Schrödinger est cherchée sous la forme :

$$\Psi(x,\, t) = \mathrm{e}^{\mathrm{i}k(t)x}\, u_{k(t)}(x,\, t) \ . \tag{30.354}$$

Par un choix judicieux de la fonction $k(t)$ (le donner), montrer que $u_{k(t)}(x,\, t)$ satisfait :

$$\mathrm{i}\hbar\frac{\partial}{\partial t}u_{k(t)}(x,\, t) = H(x,\, p\,;\, k(t))\, u_{k(t)}(x,\, t) \ . \tag{30.355}$$

2. On introduit le changement de variables suivant :

$$x' = x - \frac{1}{2}gt^2 \ , \qquad p' = p - Mgt \ . \tag{30.356}$$

Interpréter classiquement cette transformation.

3. Au changement de repère ci-dessus, correspond une transformation unitaire $\mathcal{S}$ dans l'espace des états, reliant les deux vecteurs dans un repère et un autre :

$$|\Psi'\rangle = \mathcal{S}|\Psi\rangle \iff \Psi'(x,\, t) = \mathcal{S}(x,\, p,\, t)\Psi(x,\, t) \ . \tag{30.357}$$

$|\Psi\rangle$ satisfaisant :

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\Psi\rangle = H|\Psi\rangle \ , \tag{30.358}$$

montrer que $|\Psi'\rangle$ satisfait une équation du même type avec un Hamiltonien H' dont on donnera l'expression en fonction de H, de $\mathcal{S}$ et de sa dérivée.

4. On pose :

$$\mathcal{S}(x,\, p,\, t) = \mathrm{e}^{\mathrm{i}\gamma(t)}\,\mathrm{e}^{\mathrm{i}\alpha(t)p}\,\mathrm{e}^{\mathrm{i}\beta(t)x} \ . \tag{30.359}$$

Donner les expressions des observables transformées $\mathcal{S}x\mathcal{S}^\dagger$ et $\mathcal{S}p\mathcal{S}^\dagger$.

5. En déduire l'expression de H'.

Mise en évidence de l'oscillation de la vitesse d'un atome

Le champ de nombre d'onde k_0 contient maintenant, entre les instants $t = 0$ et $t = \Delta t$, une dérive de phase $\phi(t)$ quadratique en temps ; très précisément, le Hamiltonien de l'atome est alors :

$$H_0 = \frac{p^2}{2M} + V_0 \sin^2 \left[k_0\left(x + \frac{1}{2}gt^2\right)\right] \ . \tag{30.360}$$

1. Effectuer sur cet Hamiltonien la transformation $\mathcal{S}$ introduite en (30.357), et écrire l'expression de H_0'. Trouver successivement $\alpha(t)$, puis $\beta(t)$ et enfin $\gamma(t)$ pour que H_0' se réduise à un Hamiltonien statique du genre (30.353).

2. D'après la question 1 de la partie précédente, les fonctions propres (dans le repère accéléré) sont de la forme :

$$\mathrm{e}^{\mathrm{i}k(t)x} u_{n\, k(t)}(x,\, t) \ , \tag{30.361}$$

   où $u_{n\,k(t)}$ est propre de $H(x,\, p,\, ;\, k(t))$. En adoptant l'hypothèse d'adiabaticité et sachant que l'atome est au départ dans l'état d'impulsion bien déterminée $k = 0$, quelle est sa fonction d'onde $\Psi'(x,\, t)$ à l'instant t ?

3. Quelle est la relation entre la valeur moyenne de la vitesse $v = \langle p/M \rangle$ et la valeur propre $\varepsilon(k(t))$?

4. Sans calcul (en utilisant notamment le schéma de zones répétées), montrer que la vitesse v a une variation périodique ; exprimer sa période T_{B} en fonction de $\hbar$, k_0, M et g.

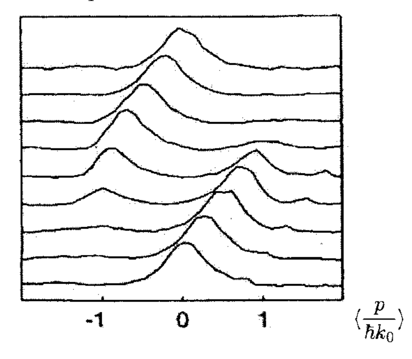


Figure 30.19: Population atomique en fonction de la vitesse réduite $\langle p/(\hbar k_0)\rangle$, pour différentes valeurs de Δt (figure extraite de [249]).

5. La figure 30.19 représente, dans le repère accéléré, la variation de la population atomique en fonction de la vitesse v, les atomes étant préparés dans un état où k est en moyenne nul et peu dispersé à l'échelle de la première zone de Brillouin .

Chaque courbe correspond à une valeur donnée de l'intervalle de temps Δt pendant lequel on applique la dérive de phase ϕ définie plus haut.

Commenter sommairement cette figure. Sachant que les courbes successives correspondent, de bas en haut, à $\Delta t = 0,\ 1, 2, \ldots$ ms, quelle est la durée d'une oscillation de Bloch ?

30.7.14 Atomes dans un réseau optique

À l'aide de trois lasers mutuellement perpendiculaires accordés à la longueur d'onde λ, on peut réaliser un réseau optique de Bravais, caractérisé par un potentiel périodique V. La maille primitive a une dimension de l'ordre d'une longueur d'onde optique. Typiquement, le potentiel V est de la forme :

$$V(x,\, y,\, z) = V_x \sin^2 kx + V_y \sin^2 ky + V_z \sin^2 kz \ , \tag{30.362}$$

avec $k = \frac{2\pi}{\lambda}$. Le réseau étant constitué, on y injecte un ensemble de N atomes dont il s'agit d'étudier la dynamique quantique. La profondeur des puits où peuvent se piéger les atomes est ajustable en faisant varier l'intensité des lasers.

Un tel système se prête bien à l'étude de questions fondamentales initialement apparues en Matière condensée, d'une part en exploitant une gamme de paramètres physiques inaccessibles dans les solides ordinaires (la taille de la maille, par exemple), d'autre part en ayant la possibilité de faire varier presque *ad libitum* des caractéristiques autrement imposées (la profondeur V_u des puits par exemple), ou en tout cas faiblement variables. Enfin la nature des atomes (bosons ou fermions) permet de mettre en évidence les différences fondamentales de comportement des deux classes de particules.

Une caractéristique physique importante du système est la densité d'atomes relativement à la dimension de la maille primitive ; on se place dans la limite où chaque site est essentiellement occupé par un petit nombre d'atomes, 0, 1 ou 2.

Dans toute la suite, l'espace des états de chaque atome i de masse M est réduit à un multiplet de spin $S_i = 1$ qui, en pratique, est une variété hyperfine de l'état fondamental de dimension $2S_i + 1$.

Interaction effective entre deux atomes

On admet que deux atomes n'interagissent que s'ils se trouvent sur le même site du réseau, et que cette interaction est essentiellement fixée par la valeur S du spin total $\vec{S} = \vec{S}_1 + \vec{S}_2$. On écrit ainsi l'interaction sous la forme traditionnelle d'une interaction de contact de Fermi[97] :

$$V_{\text{int}}(\vec{r}_1,\, \vec{r}_2) = \frac{4\pi\hbar^2}{M}\, \delta(\vec{r}_1 - \vec{r}_2) \sum_S a_S P_S \ ; \tag{30.363}$$

[97] du même genre que pour la diffusion neutron - noyau, voir chapitre (29), éq. (29.46). Le facteur 2 additionnel provient de la masse réduite $M/2$.

a_S est appelée *longueur de diffusion* ; P_S est le projecteur $\sum_{M_S} |S\,M_S\rangle\langle S\,M_S|$ dans le sous-espace associé à la valeur S du spin total.

1. Quelles sont *a priori* les valeurs possibles de S dans l'expression (30.363) ?

2. On désigne par $|\alpha\rangle$, $(\alpha = +,\, 0,\, -)$ les trois états de spin d'un atome $M_{S_i} = +1,\, 0,$ -1 et, pour simplifier, on notera $|\alpha\rangle|\alpha'\rangle$ l'état produit tensoriel $|\alpha\rangle \otimes |\alpha'\rangle$. Donner l'expression de l'état $|S = 2,\, M_S = 2\rangle \equiv |2,\, 2\rangle$ du spin total en fonction des $|\alpha\rangle|\alpha'\rangle$.

3. Par des applications successives de l'opérateur S_- en déduire :

   (a) les états $|2,\, 1\rangle$, $|2,\, 0\rangle$,

   (b) les états $|1,\, 1\rangle$, $|1,\, 0\rangle$,

   (c) l'état $|0,\, 0\rangle$.

   en fonction des $|\alpha\rangle|\alpha'\rangle$. Préciser à chaque fois la symétrie de permutation de ces états de spin.

4. Lorsqu'ils sont sur le même site, deux atomes sont forcément dans la même orbitale. En déduire que la valeur $S = 1$ est alors en fait interdite et que, par conséquent, la somme sur S dans (30.363) se réduit à deux termes.

5. À partir de $\vec{S}_1.\vec{S}_2 = \frac{1}{2}(\vec{S}^2 - \vec{S}_1^2 - \vec{S}_2^2)$, et en notant que $\vec{S}^2$ est un opérateur scalaire dans un sous-espace S fixé, montrer que :

$$\vec{S}_1.\vec{S}_2 = \hbar^2(P_2 - 2P_0) \ . \tag{30.364}$$

6. En remarquant que $P_0 + P_2 = \mathbf{1}$, montrer que :

$$V_{\text{int}}(\vec{r}_1,\, \vec{r}_2) = \frac{4\pi}{3M}\, \delta(\vec{r}_1 - \vec{r}_2)(l_0\hbar^2\mathbf{1} + l\, \vec{S}_1.\vec{S}_2) \ , \tag{30.365}$$

   où l_0 et l sont des longueurs dont on donnera l'expression en fonction des longueurs de diffusion.

7. Pour l'atome ^{23}Na, on a $a_0 \simeq 46a_{\text{B}}$, $a_2 \simeq 52a_{\text{B}}$, a_{B} désignant le rayon de la première orbite de Bohr. En déduire la nature, ferromagnétique ou antiferromagnétique, de l'interaction effective entre deux atomes de sodium situés sur le même site du réseau.

8. Soit $\phi(\vec{r})$ l'orbitale utilisée sur chaque site. Donner l'expression de U_S, énergie d'une paire d'atomes (immobiles) situés sur le même site, et en déduire l'inégalité entre U_0 et U_2.

Hamiltonien effectif en présence de passage tunnel (2 atomes)

Tout atome situé en un site du réseau a la possibilité de passer par effet tunnel vers les premiers voisins. L'objectif de cette partie est, sous l'hypothèse que le *tunneling* est faible, d'établir l'expression suivante pour le Hamiltonien effectif d'une paire d'atomes :

$$H_{\text{eff}} = \varepsilon - \hbar^{-2} J \vec{S}_1.\vec{S}_2 - \hbar^{-4} K (\vec{S}_1.\vec{S}_2)^2 \ , \qquad (30.366)$$

où les constantes ε, J et K s'expriment à l'aide des énergies U_S et de l'amplitude tunnel t (définie plus précisément ci-dessous). En temps utile, on s'appuiera sur l'hypothèse $|\frac{t}{U_S}| \ll 1$, bien vérifiée pour $a_S \ll \lambda$.

Pour deux sites premiers voisins A et B, on note ϕ_a et ϕ_b les deux orbitales localisées orthogonales, respectivement centrées sur A et sur B (chacune est la jumelle translatée de l'autre).

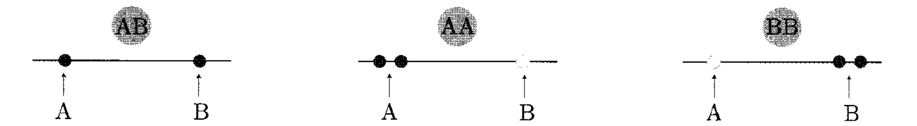


Figure 30.20: Les trois configurations considérées dans le texte : pour deux sites adjacents A et B, un atome sur chaque site, ou deux atomes sur un même site. • représente un atome, ∘ un site vide. Pour la configuration AB, le spin total peut prendre les trois valeurs $0, 1$ et 2 ; pour les configurations AA et BB, les seules valeurs possibles sont $S = 0,\ 2$.

Pour deux atomes, on écrit le Hamiltonien sous la forme :

$$H_{2\text{at}} = H_0(1,\ 2) + T(1) + T(2) \ \equiv\ H_0 + W \ . \qquad (30.367)$$

H_0 correspond aux atomes immobiles et a donc pour énergies propres 0 ($S = 0,\ 1,\ 2$) et U_S ($S = 0,\ 2$). T est l'opérateur indépendant du spin qui assure le passage d'un atome d'un site vers un premier voisin ; on pose très précisément :

$$\langle \phi_a | T | \phi_b \rangle = t \ \in \mathbb{R} \ , \qquad \langle \phi_a | T | \phi_a \rangle = \langle \phi_b | T | \phi_b \rangle = 0 \ . \qquad (30.368)$$

Modifications au second ordre dues à l'effet - tunnel

1. Pour $S = 0,\ 2$, on note $\Phi_{ij}(1,\ 2)|S,\ M_S\rangle$ ($i = a,\ b,\ j = a,\ b$) les états à deux atomes dont la partie d'espace normalisée Φ_{ij} est construite avec les orbitales ϕ_a et ϕ_b. Donner les expressions normalisées de Φ_{aa}, Φ_{ab} et Φ_{bb}.

2. Que peut-on dire des éléments de matrice de W entre deux états $|S,\ M_S\rangle$ différents ?

3. Montrer que :

$$\langle \Phi_{aa}|W|\Phi_{aa}\rangle = \langle \Phi_{bb}|W|\Phi_{bb}\rangle = \langle \Phi_{ab}|W|\Phi_{ab}\rangle = 0 \ , \tag{30.369}$$

$$\langle \Phi_{ab}|W|\Phi_{aa}\rangle = \langle \Phi_{ab}|W|\Phi_{bb}\rangle = \sqrt{2}\,t \ . \tag{30.370}$$

4. S et M_S étant fixés, écrire la matrice de $H_{2\mathrm{at}}$ sur la base ordonnée Φ_{ab}, Φ_{aa}, Φ_{bb} et son équation caractéristique.

5. En déduire les trois valeurs propres de $H_{2\mathrm{at}}$ développées jusqu'à l'ordre 2 inclus en $\frac{t}{U_S}$.

On a vu ci-dessus que pour un spin total $S = 1$, les configurations AA et BB sont interdites ; en outre, un petit calcul montre que l'élément de matrice diagonal de $H_{2\mathrm{at}}$ est nul. Au total, pour $S = 1$, le *tunnelling* n'a pas d'effet, l'énergie propre restant nulle même en présence d'effet - tunnel.

Hamiltonien effectif Dans toute la suite, on ne conserve que les trois états d'énergies les plus basses, à savoir :

$$E_{S=1} = 0 \ , \qquad E_{S=0,\,2} = -\frac{4t^2}{U_S} \ . \tag{30.371}$$

Il s'agit de montrer que l'on peut alors écrire un Hamiltonien effectif tel que H_{eff} donné en (30.366).

1. Quelles sont, pour chaque valeur de S, les valeurs moyennes de $\vec{S}_1.\vec{S}_2$ et de $(\vec{S}_1.\vec{S}_2)^2$ dans un multiplet $\{|S,\, M_S\rangle\}_{-S\leq M_S\leq +S}$ du spin total ?

2. En donnant à S ses trois valeurs possibles et en utilisant l'expression (30.366) de H_{eff}, écrire le système linéaire reliant les trois paramètres ε, J et K aux trois énergies E_S ($S = 0,\, 1,\, 2$).

3. Trouver l'expression de ces paramètres en fonction des deux énergies $E_{S=0,\,2}$. Noter que pour ^{23}Na :

$$0 < J \lesssim K \ . \tag{30.372}$$

Ceci achève la détermination du Hamiltonien effectif H_{eff} donné en (30.366) pour deux atomes.

Étude de l'état fondamental d'un réseau unidimensionnel

Dans toute la suite, on omet la constante additive ε. Par ailleurs, en supposant une forte anisotropie du potentiel, $V_x \ll V_y$, V_z, les atomes sont pratiquement confinés le long de l'axe Ox et le réseau optique est de fait un réseau unidimensionnel, dont chaque site est repéré par l'indice $n \in \mathbb{Z}$. Enfin, on se met strictement dans l'hypothèse d'un demi-remplissage, où chaque site contient 1 atome en moyenne.

En raison de ce qui précède, il est naturel de poser, pour le réseau 1-d, le Hamiltonien effectif suivant ($S_n = 1 \,\forall\, n$) :

$$H_{\text{res}} = -\hbar^{-2}\sum_n J\vec{S}_n.\vec{S}_{n+1} + \hbar^{-4}K\sum_n(\vec{S}_n.\vec{S}_{n+1})^2 \ . \tag{30.373}$$

Dans le cas-limite $K = 0$, H_{res} se réduit au Hamiltonien de Heisenberg. Le cas $K \neq 0$ a été l'objet de nombreuses études donnant lieu à des résultats parfois contradictoires.

L'une des dificultés du problème tient à l'importance de la nature quantique des spins, comme le montrent les réponses aux questions suivantes :

1. si les spins étaient de simples vecteurs au sens classique du terme, quel serait l'état fondamental de la chaîne d'atomes ? Le représenter par un petit schéma.
2. On a vu que $E_{S=0} < E_{S=2}$, de sorte que l'état fondamental d'une paire d'atomes est un état singulet. Intuitivement, comment a-t-on envie *a priori* de schématiser l'état fondamental de la chaîne ?

Dans toute la suite, on raisonne en unités J de sorte que l'opérateur d'intérêt est :

$$\tilde{H} = -\hbar^{-2}\sum_n \vec{S}_n.\vec{S}_{n+1} + \hbar^{-4}\tan\gamma\,(\vec{S}_n.\vec{S}_{n+1})^2 \ , \qquad \tan\gamma = \frac{K}{J} \ ; \tag{30.374}$$

compte tenu des valeurs numériques pour ^{23}Na, l'angle γ est dans l'intervalle $[\frac{\pi}{4},\, \frac{\pi}{2}[$.

Dans la littérature, on s'accorde à penser que si γ est voisin de $\frac{\pi}{2}$, l'état fondamental est la juxtaposition de dimères formé essentiellement de paires d'atomes (un sur chaque site) et décrit convenablement par la fonction d'onde de spin :

$$|F_{\text{dim}}\rangle = \ldots\, |F_{-3-2}\rangle\, |F_{-10}\rangle\, |F_{12}\rangle\, |F_{34}\rangle \ldots \ , \tag{30.375}$$

où $|F_{nn+1}\rangle$ est la fonction de spin total singulet construite avec deux atomes de spin $S_i = 1$ sur les sites n et $n+1$, selon la configuration AB de la fig. 30.20 étendue à un réseau 1-d. On a vu plus haut (première partie, question 2) que chacune de ces fonctions est de la forme :

$$|F_{nn+1}\rangle = \frac{1}{\sqrt{3}}(|+\rangle|-\rangle + |-\rangle|+\rangle - |0\rangle|0\rangle)_{nn+1} \ . \tag{30.376}$$

La question de la stabilité de cet état quand γ décroît à partir de $\frac{\pi}{2}$ est controversée. En particulier, un autre état, dit *nématique*, a été considéré, où chaque site contient un atome dans l'état de spin $|\alpha = 0\rangle \equiv |0\rangle$, donnant la fonction de spin :

$$|F_{\text{nem}}\rangle = \ldots\ |0\rangle_{-2}\, |0\rangle_{-1}\, |0\rangle_0\, |0\rangle_1\, |0\rangle_2 \ldots \tag{30.377}$$

Dans la suite, on examine ces questions particulières à l'aide d'une méthode variationnelle[98].

Introduction d'une fonction variationnelle et calcul de l'énergie On pose la fonction variationnelle suivante :

$$|\tilde{F}_{nn+1}\rangle = \frac{1}{\sqrt{2+\xi^2}}(|+\rangle|-\rangle + |-\rangle|+\rangle - \xi|0\rangle|0\rangle)_{nn+1} \ , \tag{30.378}$$

où ξ désigne le paramètre variationnel, supposé réel positif.

1. À quoi correspondent respectivement les deux limites $\xi = 1$ et $\xi = \infty$?
2. L'énergie variationnelle d'une paire d'atomes s'obtient en calculant la moyenne de $\tilde{H}$ (voir (30.374)) à l'aide de l'un des facteurs apparaissant dans (30.375), $\tilde{F}_{nn+1}$ par exemple. Montrer que l'énergie de la paire $(n,\, n+1)$, $\tilde{E}_{nn+1}$, est la somme de deux contributions :
$$\tilde{E}_{nn+1} = \tilde{E}^{(\text{intra})} + \tilde{E}^{(\text{inter})} \ , \tag{30.379}$$
où $\tilde{E}^{(\text{inter})}$ provient de l'interaction de la paire $(n,\, n+1)$ avec la paire $(n+1,\, n+2)$.
3. Montrer que :
$$\tilde{E}^{(\text{intra})} = \frac{2}{2+\xi^2}\left[1 + 2\xi - (3 + 2\xi + \xi^2)\tan\gamma\right] \ , \tag{30.380}$$
4. Par ailleurs, un calcul un peu long montre que :
$$\tilde{E}^{(\text{inter})} = -\frac{2}{(2+\xi^2)^2}(3 + 2\xi^2 + \xi^4)\tan\gamma \ . \tag{30.381}$$
En déduire l'expression de l'énergie variationnelle par paire $\epsilon(\xi)$.

Discussion Dans cette partie, on analyse la fonction $\epsilon(\xi)$, $\xi \geq 1$, paramétrée par γ, afin de tirer des conclusions sur la nature de l'état fondamental prévue par le traitement variationnel ci-dessus.

[98] Pour plus de détails, voir [250].

1. Un calcul un peu fastidieux montre que le développement limité de $\epsilon(\xi)$ près de $\xi = 1$ est :

$$\epsilon(\xi) = 1 - \frac{8}{3}\tan\gamma + \frac{1}{3}(\frac{5}{9}\tan\gamma - 1)(\xi - 1)^2 + \mathcal{O}((\xi - 1)^3) \ . \tag{30.382}$$

   En déduire la stabilité de l'état dimérisé selon la valeur de γ.

2. Étudier sommairement le comportement de $\epsilon(\xi)$ pour $\xi \gg 1$.

3. En déduire l'allure du graphe de $\epsilon(\xi)$ selon la valeur de γ.

4. Tirer les conclusions de l'approche variationnelle sur la nature de l'état fondamental.

30.7.15 Compétition entre confinement magnétique et interactions dans un réseau de plaquettes

Le but de ce problème est d'analyser les propriétés électroniques d'un réseau soumis à un champ magnétique $\vec{\mathcal{B}}$. Le réseau choisi est une chaîne de *plaquettes* carrées (voir fig. 30.21, à droite). Dans la première partie, on se borne à une théorie à une particule et on met en évidence un phénomène de localisation remarquable quand le flux du champ par plaquette prend certaines valeurs bien déterminées [251]. Dans la deuxième partie, on analyse l'effet de l'interaction entre deux électrons sur le phénomène de localisation mis en évidence précédemment pour un électron célibataire [252] [99].

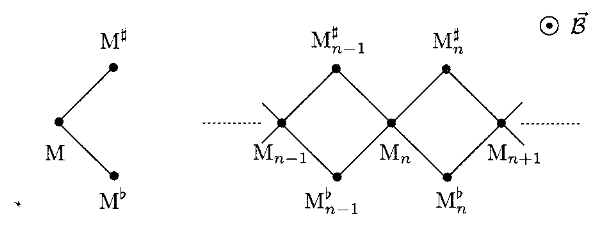


Figure 30.21: À gauche : un motif. À droite : réseau de plaquettes obtenu par juxtaposition des motifs.

Le site M_n a pour abscisse na ($n \in \mathbb{Z}$). Sur chaque site, on introduit une même fonction atomique ϕ, normalisée et de faible extension spatiale ; on note ϕ_n, $\phi_n^\sharp$ et $\phi_n^\flat$ les fonctions centrées respectivement en $\mathrm{M}_n^\sharp$, M_n et $\mathrm{M}_n^\flat$ (toutes ces fonctions se déduisent donc les unes des autres par des translations). On négligera tous les produits scalaires entre deux fonctions atomiques distinctes.

[99] Des coquilles présentes dans cet article ont été rectifiées dans la rédaction de ce problème.

États et dynamique d'un électron sur le réseau

Étude d'un motif Pour simplifier, on note ϕ, $\phi^\sharp$ et $\phi^\flat$ les trois orbitales définies sur les trois atomes d'un motif unique. Dans cette sous-section, H_m désigne le Hamiltonien d'un électron situé dans ce motif isolé.

1. On note $\mathcal{S}$ l'opérateur associé à la symétrie par rapport à l'axe horizontal passant par le point M.
   (a) Que vaut $[H_\mathrm{m}, \mathcal{S}]$?
   (b) Quelles sont les valeurs propres de $\mathcal{S}$?
   (c) Écrire l'expression du vecteur propre de $\mathcal{S}$, $|a\rangle$, associé à la valeur propre négative de $\mathcal{S}$.
   (d) En déduire la forme générale des deux autres vecteurs propres de $\mathcal{S}$; quelle est la valeur propre de $\mathcal{S}$ correspondante ?
   (e) En déduire que $|a\rangle$ est aussi vecteur propre de H_m.
2. Pour simplifier, on ne considère que les amplitudes de saut entre plus proches voisins, et on pose :

$$\langle\phi|H_\mathrm{m}|\phi^\sharp\rangle = \langle\phi|H_\mathrm{m}|\phi^\flat\rangle = v \qquad (v \in \mathbb{R}_+) \ . \tag{30.383}$$

   Les éléments diagonaux de H_m, tous égaux entre eux, définissent l'origine de l'énergie.
   (a) Quelles sont les valeurs propres de H_m ?
   (b) Exprimer les vecteurs propres correspondants en fonction de $|\phi\rangle$, $|\phi^\sharp\rangle$ et $|\phi^\flat\rangle$.

États électroniques du réseau en présence d'un champ magnétique On sait que pour une particule de masse m et de charge q en présence d'un champ magnétique $\vec{\mathcal{B}}$, le Lagrangien s'écrit :

$$L = \frac{1}{2}m\vec{v}^{\,2} + q\,\vec{A}.\vec{v} \ , \tag{30.384}$$

où $\vec{A}$ est le potentiel vecteur ($\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$).

1. Feynman a montré que le propagateur quantique $U(\vec{r},\, t\,;\, \vec{r}_0,\, t_0)$ peut le plus souvent[100] s'obtenir en effectuant une somme sur tous les chemins possibles de l'exponentielle de l'action, chemins allant de $(\vec{r}_0,\, t_0)$ à $(r,\, t)$. Symboliquement, ce résultat s'écrit :

$$U(\vec{r},\, t\,;\, \vec{r}_0,\, t_0) = \sum_{\text{chemins}} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} L\mathrm{d}t'} \ . \tag{30.385}$$

[100] On admettra que c'est le cas ici.

(a) Écrire très précisément le facteur additionnel[101] survenant lorsque l'on applique le champ magnétique, en y faisant apparaître une intégrale curviligne représentant la circulation du potentiel-vecteur.

(b) Ce facteur est-il invariant[102] de jauge ?

(c) Qu'en est-il si l'intégrale curviligne est prise sur un circuit fermé (boucle) ?

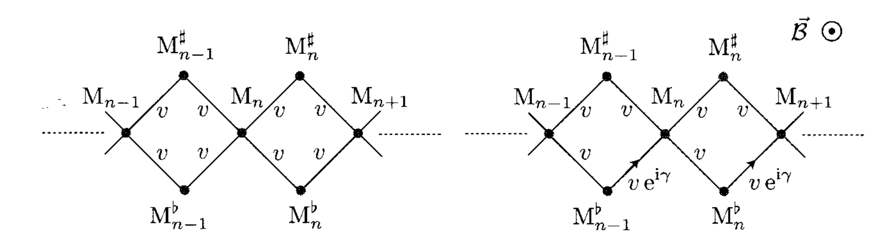


Figure 30.22: À gauche : situation en l'absence de champ magnétique. À droite : γ est la phase additionnelle due au champ magnétique, relative à l'un des quatre liens.

2. Transposés au réseau traité en liaisons fortes, où l'électron passe d'un site à l'autre, les résultats précédents se traduisent par une phase additionnelle sur les éléments de matrice du Hamiltonien entre deux fonctions localisées sur les sites constituant les extrémités du chemin. Le champ $\vec{\mathcal{B}}$ est orthogonal au plan du réseau de plaquettes.

   Par un choix judicieux de la jauge, il est possible de n'altérer qu'une seule des quatre amplitudes associées aux quatre liens d'une plaquette ; la figure 30.22 décrit précisément le choix adopté dans la suite – la flèche sur le lien de $M^\flat_n$ vers M_{n+1} exprime le fait que l'élément de matrice correspondant est maintenant complexe. H désignant le Hamiltonien d'un électron se déplaçant sur le réseau en présence du champ magnétique, on a très précisément :

$$\langle\phi_{n+1}|H|\phi^\flat_n\rangle = v\,e^{i\gamma} \ , \tag{30.386}$$

   les trois autres amplitudes de saut étant inchangées :

$$\langle\phi_n|H|\phi^\flat_n\rangle = \langle\phi_n|H|\phi^\sharp_n\rangle = \langle\phi^\sharp_n|H|\phi_{n+1}\rangle = v \ . \tag{30.387}$$

   (a) Montrer que :

$$\gamma = \frac{e\mathcal{B}a^2}{2\hbar} \equiv 2\pi\,\frac{\mathcal{B}a^2}{2(h/e)} \ , \tag{30.388}$$

   et interpréter le résultat[103].

   (b) Estimer la valeur du champ $\mathcal{B}$ pour avoir $\gamma \sim 1$ dans les deux cas suivants :

      i. le réseau est de nature atomique ($a \sim 1$ Å) ;

[101] La phase additionnelle est responsable de l'effet de Bohm-Aharonov.

[102] On rappelle qu'un changement de jauge s'exprime notamment par $\vec{A}' = \vec{A} + \vec{\nabla} f$.

[103] Le rapport $\varphi_0 \stackrel{\text{déf}}{=} \frac{h}{e}$ est appelé *flux élémentaire*.

ii. le réseau est constitué de “points quantiques” (*quantum dots*) reliés par des fils de longueur mésoscopique (a est de l'ordre du micron).

3. Il s'agit maintenant de trouver les états propres d'un électron sur le réseau.

   (a) À l'aide des vecteurs $|\phi_n\rangle$, $|\phi^\sharp_{n'}\rangle$ et $|\phi^\flat_{n''}\rangle$, former trois classes d'états de Bloch, notés $|\psi_k\rangle$, $|\psi^\sharp_k\rangle$ et $|\psi^\flat_k\rangle$ (justifier la réponse en montrant que les états ainsi construits satisfont bien le théorème de Bloch). Normaliser[104] ces états.

   (b) On désigne par $\mathcal{F}$ la transformation permettant de passer des $|\phi\rangle$ aux $|\psi\rangle$:

$$|\phi\rangle \xrightarrow{\mathcal{F}} |\psi\rangle \ . \tag{30.389}$$

   Les états $|\psi\rangle$ sont-ils localisés ou étendus ?

   (c) Calculer $H|\psi_k\rangle$, $H|\psi^\sharp_k\rangle$ et $H|\psi^\flat_k\rangle$

   (d) En déduire la matrice de H sur la base $(|\psi_k\rangle, |\psi^\sharp_k\rangle, |\psi^\flat_k\rangle)$

   (e) On pose :

$$R(ka) = 2\sqrt{1+\cos\frac{\gamma}{2}\cos ka} \ . \tag{30.390}$$

   Trouver les valeurs propres de H, que l'on notera $\varepsilon_\alpha(k)$, $\alpha = 0, \pm 1$.

   (f) Représenter la variation des $\varepsilon_\alpha(\vec{k}\,)$ dans la première zone de Brillouin. Comment se comparent-elles aux valeurs propres d'un motif isolé ?

   (g) Que se passe-t-il pour les valeurs propres à *demi-flux*, c'est-à-dire si $\gamma = \pi$?

   (h) Revenant à γ quelconque, on pose :

$$\lambda_0 = 1 + \mathrm{e}^{-ika} \ , \qquad \lambda = 1 + \mathrm{e}^{+i\gamma}\mathrm{e}^{-ika} \ . \tag{30.391}$$

   Former les vecteurs propres normalisés $|\Psi_{k,\,\alpha}\rangle$ correspondants.

   (i) Inversement, exprimer $|\psi_k\rangle$ en fonction des $|\Psi_{k,\,\pm 1}\rangle$.

Isolant vs. conducteur ? Dans cette sous-section, on considère l'ensemble des électrons apportés par chaque atome et on remplit les bandes conformément à la prescription habituelle pour former l'état fondamental du cristal en l'absence d'interactions entre électrons.

1. Les bandes $\varepsilon_\alpha(k)$ sont-elles toutes dispersées ?

2. Quelle est, en fonction de γ, la valeur du *gap* en bord de zone ?

3. Quelle est la vitesse (moyenne) $v_\alpha(\vec{k}\,)$ d'un électron de Bloch dans chacune de ces bandes ?

[104] Ici comme ailleurs, on fera référence à des conditions cycliques de Born - von Kármán définies sur $N \gg 1$ mailles élémentaires.

4. Soit respectivement n_e, $n_\text{e}^\sharp$ et $n_\text{e}^\flat$ les nombres d'électrons apportés par chaque atome situé en M_n, $\text{M}_n^\sharp$ et $\text{M}_n^\flat$; discuter brièvement les propriétés de conduction à température nulle du réseau dans les cas suivants :

   (a) $n_\text{e} = n_\text{e}^\sharp = n_\text{e}^\flat = 1$;

   (b) $n_\text{e} = 3$, $n_\text{e}^\sharp = n_\text{e}^\flat = 1$.

5. Si $n_\text{e} = 2$, $n_\text{e}^\sharp = n_\text{e}^\flat = 1$, comment varie schématiquement en fonction de $\mathcal{B}$ la conductivité à température donnée non-nulle[105] ?

Dynamique d'un paquet d'ondes On suppose l'électron initialement situé au site $\text{M}_{n=0}$, dans l'état $|\phi_{n=0}\rangle$:

$$|\Psi(t=0)\rangle = |\phi_0\rangle \ ; \tag{30.392}$$

il s'agit de trouver l'état à un instant ultérieur, $|\Psi(t)\rangle$, et d'analyser le mouvement spatial de l'électron ainsi préparé.

1. En raisonnant intuitivement, décrire qualitativement le mouvement attendu du paquet d'ondes.
2. Exprimer $|\phi_n\rangle$ en fonction des $|\psi_k\rangle$.
3. En déduire $|\phi_{n=0}\rangle$ en fonction des $|\psi_k\rangle$, puis en fonction des $|\Psi_{k,\,\alpha}\rangle$.
4. Afin de calculer commodément l'évolution, on écrit la décomposition spectrale de l'opérateur d'évolution $\mathcal{U}(t)$:

$$\mathcal{U}(t) = \sum_k |\Psi_{k,\,0}\rangle\langle\Psi_{k,\,0}| + \dots \ ; \tag{30.393}$$

   achever d'écrire le second membre.

5. En déduire l'expression de l'état $|\Psi(t)\rangle$ développé sur les $|\psi_k\rangle$, $|\psi_k^\sharp\rangle$ et $|\psi_k^\flat\rangle$.
6. Soit A_{00} l'amplitude $\langle\phi_0|\Psi(t)\rangle$. Que représente la quantité $P_0(t) \stackrel{\text{déf}}{=} |A_{00}(t)|^2$?
7. Montrer que dans la limite $N \to \infty$:

$$A_{00}(t) = \frac{1}{\pi}\int_0^{+\pi} \cos\left[\frac{vt}{\hbar}R(x)\right] \mathrm{d}x \ , \tag{30.394}$$

   où la fonction R est définie en (30.390).

8. Comment varie $P_0(t)$ aux petits temps ? Pourrait-on *a priori* observer l'effet Zénon ?

[105] C'est un exemple d'oscillation de la magnétorésistance.

9. L'amplitude $A_{00}(t)$ s'écrit aussi :

$$A_{00}(t) = \Re I(\tau) \ , \quad I(\tau) = \frac{1}{\pi}\int_0^{+\pi} \mathrm{e}^{\mathrm{i}\tau R(x)}\,\mathrm{d}x \ , \qquad \tau = \frac{vt}{\hbar} \ . \tag{30.395}$$

À l'aide d'un argument de phase stationnaire[106] on peut montrer que, pour $\tau \gg 1$:

$$\int_0^{+\pi} \mathrm{e}^{\mathrm{i}\tau R(x)}\,\mathrm{d}x \simeq \eta\sqrt{\frac{2\pi}{\tau}}\sum_j \frac{1}{\sqrt{-\mathrm{i}R''(x_j)}}\,\mathrm{e}^{\mathrm{i}\tau R(x_j)} \ ; \tag{30.396}$$

x_j désigne un *extremum* de $R(x)$ et η vaut 1 ou $\frac{1}{2}$ selon que x_j est dans l'intervalle $]0,\,\pi[$ ou à une extrémité. Le symbole $\sqrt{z}$ désigne la branche de la racine carrée qui prend des valeurs réelles positives sur le demi-axe réel positif.

En déduire que pour $t \gg \frac{\hbar}{v}$ et si $0 \leq \gamma < \pi$:

$$P_0(t) \simeq \frac{1}{4\pi\tau\cos(\gamma/2)}\Big\{\sum_{j=0,\,1}\sqrt{R_j}\cos\Big[\frac{vt}{\hbar}R_j - (-1)^j\frac{\pi}{4}\Big]\Big\}^2 \ , \tag{30.397}$$

où $R_j \stackrel{\text{déf}}{=} 2\sqrt{1+(-1)^j\cos\frac{\gamma}{2}}$. Commenter ce résultat. Dans quelle mesure confirme-t-il la prévision intuitive de la question 1 ci-dessus ?

10. L'expression (30.397) n'a pas de sens pour $\gamma = \pi$ (demi-flux), une valeur pour laquelle un phénomène remarquable se produit.

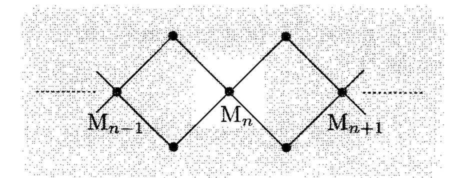


Figure 30.23: Vue d'artiste (?!) de la cage de confinement d'un électron pour le cas de demi-flux ($\gamma = \pi$).

(a) Trouver $P_0(t)$ pour $\gamma = \pi$.

(b) En déduire qu'à demi-flux l'électron reste *confiné*, tout en *respirant*.

(c) Écrire explicitement l'expression de $|\Psi(t)\rangle$.

(d) En déduire que l'électron oscille entre le site M_0 et la couronne formée par ses plus proches voisins : on peut ainsi dire que[107] *l'électron est enfermé dans une cage.*

---

[106] Il s'agit plus généralement d'une méthode de col (*saddle-point method*).

À titre de divertissement (!?), il est vivement suggéré d'une part d'établir la formule (30.396), d'autre part de comparer numériquement l'expression (30.397) au résultat exact obtenu par quadrature (numérique) de l'expression (30.394).

[107] Un tel domaine réservé est appelé *cage de Aharonov-Bohm.*

(e) Que représente la figure 30.24 ? Commenter.

(f) Soit $\mathcal{A}_0$ l'amplitude de probabilité en champ nul pour aller d'un atome à l'un de ses plus proches voisins. Expliquer le confinement à demi-flux en combinant les amplitudes des chemins possibles conduisant de M_n à M_{n+1}.

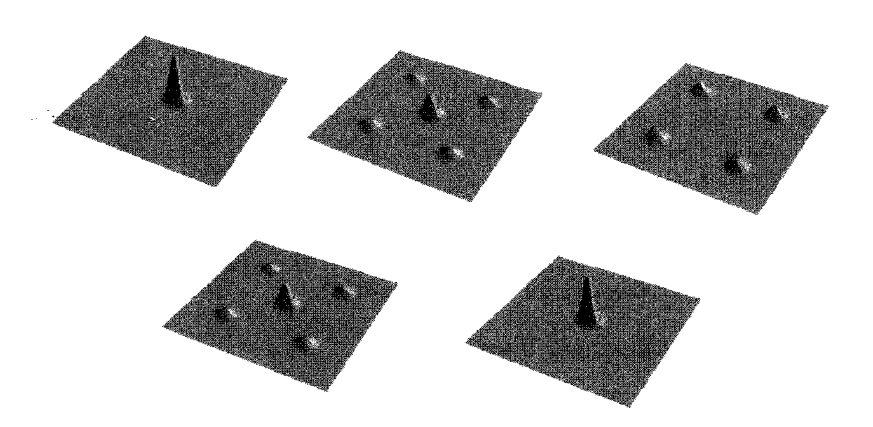


Figure 30.24: Que représentent ces dessins ?

Effet des interactions entre deux électrons

Les derniers résultats montrent qu'il existe un confinement – surprenant – aux valeurs demi-entières du flux réduit. Un point important est de savoir si cet effet de *localisation* est robuste et résiste à d'autres effets inévitables, comme l'interaction entre électrons. Dans toute cette section, on suppose que deux électrons sont déposés sur le réseau et interagissent entre eux.

Afin de modéliser le plus simplement possible leur répulsion, on choisit une interaction *à la* Hubbard : si les deux électrons ont des spins opposés *et* s'ils sont sur le même site, ils se repoussent avec l'énergie $U > 0$; dans tous les autres cas, les deux électrons ne se voient pas (classiquement parlant). Formellement, le Hamiltonien des deux électrons s'écrit :

$$\mathcal{H} = H_1 + H_2 + H_{\text{int}} \tag{30.398}$$

Propriétés générales du problème en interaction Il s'agit de discuter sommairement le cadre général du problème des deux électrons en interaction, avec très peu de calculs.

1. Commenter brièvement ce choix pour H_{int}.

2. Dans cette question, on ignore le spin.

   (a) Soit $|\Psi_{k,\alpha}\rangle$ ($\alpha = 0, \pm 1$) l'un quelconque des états obtenus dans la théorie à une particule (voir première partie) ; $\mathcal{T}_n$ désignant l'opérateur de translation tel que $\mathcal{T}_n f(x) = f(x - na)$, $\forall\, f(x)$, combien vaut $\mathcal{T}_n|\Psi_{k,\alpha}\rangle$?

   (b) Soit l'état à deux particules $|\Psi_{k,\alpha}\rangle \otimes |\Psi_{k',\alpha'}\rangle \equiv |k, \alpha, k', \alpha'\rangle$. Combien vaut $\mathcal{T}_n|k, \alpha, k', \alpha'\rangle$?

   (c) Le Hamiltonien $\mathcal{H}$ défini en (30.398) est-il invariant par les translations du réseau ?

   (d) Indépendamment du spin, les états à deux particules peuvent être cherchés comme des combinaisons linéaires des $|k, \alpha, k', \alpha'\rangle$. Expliquer pourquoi $\mathcal{H}$ ne couple entre eux que des états ayant la même valeur pour la somme $k+k' \equiv K$.

   (e) Quel sens physique peut-on attribuer à K ? Commenter en utilisant l'image d'une collision entre deux particules ayant des impulsions données et interagissant entre elles par un terme de contact.

3. Quelles sont les valeurs du spin total ?

4. Que se passe-t-il pour les états triplets ?

5. Avec deux orbitales quelconques f et g, on peut former les états $|f\rangle \otimes |g\rangle \equiv |f\, g\rangle$. Écrire la partie d'espace de l'état singulet construit avec ces deux orbitales (distinguer les deux cas $f \neq g$ et $f = g$). Un tel état est noté généralement $|f\, g\rangle_{\text{S}}$ dans la suite.

Étude détaillée du cas $\gamma = \pi$ La question étant de savoir si la localisation spectaculaire à demi-flux persiste en présence d'interaction, on choisit strictement $\gamma = \pi$ dans toute la suite. Dans cette situation particulière, chaque électron préparé dans sa cage a envie d'y rester confiné, mais la répulsion de l'autre a plutôt tendance à l'en expulser : le point est donc de voir si la répulsion rend la localisation impossible ou si celle-ci est assez robuste pour se maintenir malgré tout. Techniquement, ce qu'il faut établir est l'existence ou l'inexistence, en présence de l'interaction, d'une dispersion pour les énergies d'une paire d'électrons.

1. Écrire précisément les vecteurs propres $|\Psi_{k,\alpha}\rangle$ pour $\gamma = \pi$.

2. Effectuer la transformation $\mathcal{F}^{-1}$ (voir (30.389)) sur les $|\Psi_{k,\alpha}\rangle$ pour construire des états $|\Phi_{n,\alpha}\rangle$:

$$|\Psi_{k,\alpha}\rangle \stackrel{\mathcal{F}^{-1}}{\longrightarrow} |\Phi_{n,\alpha}\rangle \ . \tag{30.399}$$

   (a) Quelle est la nature des états $|\Phi_{n,\alpha}\rangle$?

   (b) Faire des figures d'un fragment de réseau où sont portés (à un facteur près) les coefficients des $|\Phi_{n,\alpha}\rangle$ développés sur les $|\phi_n\rangle$, $|\phi_n^\sharp\rangle$ et $|\phi_n^\flat\rangle$.

3. Compte tenu de la conservation de l'impulsion totale lors d'une collision entre deux électrons, on introduit les fonctions d'espace (singulet) correspondant à une impulsion totale donnée, K :

$$|\chi_L(K,\,\alpha,\,\alpha')\rangle = \frac{1}{\sqrt{N}}\sum_{n\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}Kna}|\Phi_{n,\,\alpha}\Phi_{n+L,\,\alpha'}\rangle_{\mathrm{S}} \tag{30.400}$$

$L \in \mathbb{Z}$ désigne la distance en unités a des deux électrons dans l'état $|\chi_L(K,\,\alpha,\,\alpha')\rangle$. Clairement, si $L \geq 2$, l'interaction de contact choisie est sans effet et les états $|\chi_L\rangle$ sont propres de $\mathcal{H}$, avec une énergie propre qui est la simple somme $\varepsilon_\alpha + \varepsilon_{\alpha'}$.

Seuls importent donc les états $L = 0,\,1$. On peut en fait montrer que $\mathcal{H}$ ne couple pas[108] $|\chi_{L\,=\,0}\rangle$ et $|\chi_{L\,=\,1}\rangle$, en raison des phases particulières qui apparaissent dans la définition des $|\Phi_{n,\,\alpha}\rangle$. Dans la suite, on se borne à étudier le mouvement d'une paire "liée" $L = 0$: la question posée se réduit à la possibilité de la cohabitation de deux électrons au sein d'une même cage en présence d'interaction.

(a) K étant fixé, quelle est la dimension du sous-espace engendré par les vecteurs $|\chi_0(K,\,\alpha,\,\alpha')\rangle$ *a priori* couplés entre eux par $\mathcal{H}$?

(b) Un calcul laborieux mais sans aucune difficulté permet d'établir le polynôme caractéristique fixant les valeurs propres ε. Elles sont données par :

$$\begin{aligned} P(\varepsilon\,; v,\,U,\,K) \stackrel{\text{déf}}{=} \varepsilon^6 - 2U\varepsilon^5 + (U^2 - 20v^2)\varepsilon^4 + 28v^2U\varepsilon^3 +\\ 8v^2(8v^2 - U^2)\varepsilon^2 - 8v^4U(7 + 3\cos Ka)\varepsilon +\\ +8v^4U^2(1+\cos Ka) = 0 \end{aligned} \tag{30.401}$$

A priori, que peut-on dire à propos de la dispersion des valeurs propres ?

(c) Toutefois, l'existence de solutions non dispersées n'est pas totalement exclue, et c'est le point que l'on examine maintenant.

   i. À quelle condition sur ε le polynôme P ne dépend-il pas de K ?

   ii. En déduire la condition *formelle* sur v et U pour qu'il existe une valeur propre non dispersée[109].

(d) Il s'agit maintenant de trouver les valeurs propres approchées dans les deux limites $U \ll v$ et $v \ll U$.

   i. $\sigma \stackrel{\text{déf}}{=} \frac{U}{v} \ll 1$

      A. À la limite $U \to 0$, quelles sont les valeurs propres $\varepsilon^{(0)}$ et leur dégénérescence ?

      B. Un comptage rapide des puissances dans (30.401) indique que la première correction relative est en σ. Poser $\varepsilon = \varepsilon^{(0)}(1+c\sigma)$ et en déduire que les valeurs propres sont, à cet ordre, à peu près égales à :

$$\mathcal{O}(\frac{U^2}{v})\ ,\quad \pm 2v \pm \frac{1}{4}(1-\cos Ka)U\ ,\quad \pm 4v \pm \frac{1}{16}(5+\cos Ka)U\ , \tag{30.402}$$

      où les signes sont solidaires les uns des autres

[108] Au total, on peut dire que "L est un bon nombre quantique".

[109] Cette dernière condition est satisfaite pour $U = 6\,v$.

ii. $\rho \stackrel{\text{déf}}{=} \frac{v}{U} \ll 1$

A. À la limite $v \to 0$, quelles sont les valeurs propres $\varepsilon^{(\infty)}$ et leur dégénérescence ?

B. Poser $\varepsilon = \varepsilon^{(\infty)}(1 + C\rho)$ et en déduire que les six valeurs propres sont à peu près égales[110] à :

$$U + 2(3 \pm \sqrt{5 + 4\cos Ka})\,\frac{v^2}{U}\ , \qquad \pm 2v\sqrt{1 \pm |\sin\frac{Ka}{2}|}\ , \qquad (30.403)$$

où les signes sont indépendants les uns des autres.

(e) Que peut-on généralement conclure en ce qui concerne le confinement à demi-flux quand les deux électrons sont en interaction ?

[110] Noter que, dans cette limite, quatre d'entre elles ne dépendent pas de U.

physique

Chapitre 31

Vibrations d'un solide ordonné

Ce chapitre est une introduction aux vibrations d'un réseau. Après avoir défini les branches acoustique et optique, la quantification du champ des déplacements est effectuée, conduisant tout naturellement à la notion de phonon. L'implication des phonons dans les propriétés thermodynamiques à basse température est illustrée par leur contribution à la chaleur spécifique, fournissant l'explication de l'échec de la loi classique de Dulong et Petit ; l'analogie avec la loi de Planck pour le rayonnement à l'équilibre permet de saisir une fois encore l'importance cruciale de la quantification. Enfin, on revient sur la diffraction pour montrer l'effet des vibrations sur les taches de Bragg, introduire le facteur de Debye - Waller et expliquer l'importance de la dimensionnalité sur la robustesse de l'hypersélectivité.

Jusqu'à présent, les ions (atomes, cœurs, ions, etc.) ont été considérés comme fixes, situés aux nœuds d'un réseau de Bravais $\mathcal{B}$, équivalent pour un cristal de la première étape de Born et Oppenheimer pour une molécule (cœurs totalement inertes). Ce cadre restrictif doit être dépassé, pour plusieurs raisons, dont les plus importantes sont les suivantes :

- les vibrations du réseau mettent en jeu des degrés de liberté intervenant dans la chaleur spécifique c_V. Leur quantification permet de rendre compte du comportement basse température de celle-ci, l'un des premiers succès de la quantification (Einstein, 1905), dans le droit fil de l'idée-maîtresse de Planck (1900) : contrairement à ce qu'affirme la loi (classique) de Dulong et Petit, c_V n'est pas constante mais chute à zéro à basse température, exhibant notamment, pour un isolant, un régime[1] $\propto T^3$. Tout naturellement, ces comportements non classiques se manifes-

[1]Comme on l'a vu (discussion à propos des magnons, chapitre 30, section 30.6, et problème 30.7.7), ce comportement en loi-puissance est typique de l'implication de relations de dispersion sans *gap* ayant le même type de variation. Dans son modèle, Einstein avait trouvé une chute exponentielle de c_V.

tent lorsque $k_{\text{B}}T \lesssim \hbar\omega_{\text{vibr}}$, où ω_{vibr} est une pulsation typique des vibrations du réseau.

- La conduction dans un métal relève usuellement de la loi d'Ohm, laquelle ne peut être comprise sans l'intervention d'une dissipation d'énergie (voir chapitre 32, section 32.1). Les vibrations y jouent un rôle majeur : le mouvement des électrons est un peu un jeu de billard où ceux-ci effectuent une sorte de marche au hasard en raison des collisions sur les ions (cœurs) du réseau. La dissipation est assurée par le développement d'un processus de nature aléatoire impliquant des collisions *inélastiques*[2], rendues toujours possibles par la mise en mouvement des cœurs.
- les spectres de diffraction – dont la théorie a été faite comme si les centres diffuseurs étaient fixes – sont inévitablement altérés par les vibrations. Leur analyse fine est source de précieuses informations. En outre, il s'agit d'un exemple de problème où la dimensionnalité du réseau joue un rôle important.

La question des vibrations des centres massifs se pose en soi, indépendamment de l'arrangement spatial de ces derniers (cristallisé, vitreux, amorphe, etc.). On ne considère dans la suite que les vibrations d'un cristal parfait, complètement défini par la donnée de sa structure (réseau de Bravais et motif). C'est le même point de départ que pour la description théorique des spectres IR et microondes d'une molécule, où une configuration d'équilibre du squelette est donnée, définie par la minimisation de l'énergie électronique par rapport aux positions des noyaux. L'équivalent pour un cristal est la minimisation de l'énergie de cohésion, comme exposé dans le chapitre 29. Dans toute la suite, la dynamique des électrons ayant été moyennée, seules subsistent les positions des cœurs en tant que variables dynamiques ; susceptibles de faire l'objet de la quantification au moment voulu, elles donnent lieu à la définition des excitations élémentaires du champ des vibrations (*phonons*).

31.1 L'approximation harmonique

Tout comme pour l'étude des mouvements du squelette moléculaire, et pour la même raison de fond, l'analyse des vibrations d'un solide (ordonné) repose sur l'idée que le cristal existe, et notamment que l'énergie (cinétique et potentielle) due à la vibration reste petite par rapport à l'énergie de cohésion (dans le cas contraire, cela signifierait que le cristal est proche de son point de fusion). Ceci introduit automatiquement une hypothèse de petites oscillations, l'amplitude des vibrations devant rester faible comparée aux paramètres de maille[3]. Dans toute la suite, on maintient donc l'hypothèse de l'existence d'un réseau de Bravais de référence – les nœuds de ce réseau sont alors les

[2] Ceci ne préjuge en rien de l'existence d'autres causes de dissipation : impuretés chimiques ou structurelles, dislocations,...

[3] Dans l'hypothèse harmonique qui va être faite, les énergies cinétique et potentielle de vibration sont égales entre elles, chacune représentant très exactement la moitié de l'énergie totale de vibration. On utilise parfois le mot *élastique* au lieu de *harmonique*, par opposition à *plastique* ; dans ce dernier cas, une fois la contrainte relâchée, le solide ne revient pas à son point de départ.

positions *moyennes* (d'équilibre) des objets qui vibrent – autour desquelles les cœurs maintenant mobiles effectuent des petits mouvements. Notant $\vec{r}(\vec{R})$ la position instantanée de l'ion[4] vibrant près du nœud $\vec{R}$, on peut toujours écrire :

$$\vec{r}(\vec{R}) = \vec{R} + \vec{u}(\vec{R}) \ , \tag{31.1}$$

$\vec{u}(\vec{R})$ désignant l'écart à la position d'équilibre $\vec{R}$. Le Hamiltonien du réseau est :

$$H = \sum_{\vec{R}\in\mathcal{B}} \frac{\vec{P}^{\,2}_{\vec{r}(\vec{R})}}{2M} + \frac{1}{2} \sum_{(\vec{R},\vec{R}')\in\mathcal{B}^2} V\big(\vec{r}(\vec{R}) - \vec{r}(\vec{R}')\big) \equiv T + U \ , \tag{31.2}$$

et, dans la limite des masses infinies (totale inertie, $\vec{r}(\vec{R}) = \vec{R}$), reproduit bien le Hamiltonien sur lequel est fondée la théorie de l'énergie de cohésion exposée dans le chapitre 29 (section 29.2). Compte tenu de (31.1), l'énergie potentielle s'écrit :

$$U = \frac{1}{2} \sum_{(\vec{R},\vec{R}')\in\mathcal{B}^2} V\big(\vec{R} - \vec{R}' + \vec{u}(\vec{R}) - \vec{u}(\vec{R}')\big) \ , \tag{31.3}$$

et, dans l'hypothèse où les écarts sont faibles, le développement de Taylor au voisinage des positions d'équilibre $\vec{R}$ et $\vec{R}'$ donne $U = U_{\text{eq}} + U_{\text{h}}$ avec[5] :

$$U_{\text{eq}} = \frac{1}{2} \sum_{(\vec{R},\vec{R}')\in\mathcal{B}^2} V(\vec{R} - \vec{R}') \ , \tag{31.4}$$

qui correspond au réseau figé, et :

$$U_{\text{h}} = \frac{1}{2} \sum_{(\vec{R},\vec{R}')\in\mathcal{B}^2} \sum_{\alpha,\,\beta=x,\,y,\,z} \frac{1}{2!}[\vec{u}(\vec{R}) - \vec{u}(\vec{R}')]_\alpha \, K_{\alpha\beta}(\vec{R} - \vec{R}') \, [\vec{u}(\vec{R}) - \vec{u}(\vec{R}')]_\beta \ . \tag{31.5}$$

Dans cette expression, les constantes $K_{\alpha\beta}$ sont les dérivées secondes :

$$K_{\alpha\beta}(\vec{R} - \vec{R}') = \Big(\frac{\partial^2 V}{\partial r_\alpha \partial r_\beta}\Big)_{\vec{R},\,\vec{R}'} \ , \tag{31.6}$$

calculées aux positions d'équilibre des cœurs ; la notation $K_{\alpha\beta}$ évoque une constante de raideur de ressort parfait ($[K_{\alpha\beta}] = \text{MT}^{-2}$). En réarrangeant les sommations dans (31.5), le terme harmonique U_{h} donné par (31.5) peut être mis sous la forme :

$$U_{\text{h}} = \frac{1}{2} \sum_{(\vec{R},\vec{R}')\in\mathcal{B}^2} \sum_{\alpha,\,\beta=x,\,y,\,z} \big[u_\alpha(\vec{R})u_\beta(\vec{R}) - u_\alpha(\vec{R})u_\beta(\vec{R}')\big] \, K_{\alpha\beta}(\vec{R} - \vec{R}') \tag{31.7}$$

ou encore sous la forme plus symétrique :

$$\boxed{U_{\text{h}} = \frac{1}{2} \sum_{(\vec{R},\vec{R}')\in\mathcal{B}^2} \sum_{\alpha,\,\beta=x,\,y,\,z} u_\alpha(\vec{R}) \, Q_{\alpha\beta}(\vec{R} - \vec{R}') \, u_\beta(\vec{R}')} \tag{31.8}$$

[4]Pour simplifier les notations, on suppose ici le réseau sans base. La généralisation, à ce stade, est un pur jeu d'écriture. Le cas d'un réseau avec base est examiné dans la suite (section 31.3).

[5]Le terme U_{eq} ne joue aucun rôle en ce qui concerne les vibrations, et sera omis dans la suite.

où la matrice $Q_{\alpha\beta}$ est définie comme :

$$Q_{\alpha\beta}(\vec{R}-\vec{R}') \stackrel{\text{déf}}{=} -K_{\alpha\beta}(\vec{R}-\vec{R}') + \delta_{\vec{R}\vec{R}'} \sum_{\vec{R}''} K_{\alpha\beta}(\vec{R}-\vec{R}'') \tag{31.9}$$

Par ailleurs, si on donne le même déplacement $\vec{u}_0$ à tous les atomes (translation en bloc), l'énergie potentielle ne change pas ; en conséquence :

$$0 = \frac{1}{2} \sum_{(\vec{R},\vec{R}')\in\mathcal{B}^2} \; \sum_{\alpha,\,\beta=x,\,y,\,z} u_{0,\,\alpha}(\vec{R})\, Q_{\alpha\beta}(\vec{R}-\vec{R}')\, u_{0,\,\beta}(\vec{R}') \;, \tag{31.10}$$

soit :

$$\sum_{\alpha,\,\beta=x,\,y,\,z} u_{0,\,\beta}(\vec{R}')u_{0,\,\alpha}(\vec{R}) \sum_{(\vec{R},\vec{R}')\in\mathcal{B}^2} Q_{\alpha\beta}(\vec{R}-\vec{R}') = 0 \qquad \forall\, \vec{u}_0(\vec{R}) \;; \tag{31.11}$$

la somme de droite est $N\sum_{\vec{R}} Q_{\alpha\beta}(\vec{R})$, d'où :

$$\boxed{\sum_{\vec{R}\in\mathcal{B}} Q_{\alpha\beta}(\vec{R}) = 0 \quad \forall\; \alpha,\, \beta} \tag{31.12}$$

Une énergie potentielle étant définie, il est possible d'écrire les équations du mouvement (classiques) pour chaque coordonnée $\vec{r}(\vec{R})$, étant entendu que $\ddot{\vec{r}}(\vec{R}) = \ddot{\vec{u}}(\vec{R})$; les équations de Newton sont :

$$M\,\ddot{u}_\alpha(\vec{R}) = -\frac{\partial}{\partial u_\alpha(\vec{R})}\,\frac{1}{2} \sum_{\vec{R}',\,\vec{R}''} \sum_{\alpha',\,\beta'} u_{\alpha'}(\vec{R}')Q_{\alpha'\beta'}(\vec{R}'-\vec{R}'')\, u_{\beta'}(\vec{R}'') \;; \tag{31.13}$$

la dérivation donne $\delta_{\alpha'\alpha}\delta_{\vec{R}'\vec{R}}\, u_{\beta'}(\vec{R}'') + \delta_{\beta'\alpha}\delta_{\vec{R}''\vec{R}}\, u_{\alpha'}(\vec{R}')$; en jouant avec le nom des indices muets de sommation, les équations du mouvement prennent la forme :

$$M\ddot{u}_\alpha(\vec{R}) = -\frac{1}{2}\sum_{\vec{R}'}\sum_{\beta}[Q_{\alpha\beta}(\vec{R}-\vec{R}') + Q_{\beta\alpha}(\vec{R}'-\vec{R})]\, u_\beta(\vec{R}') \;. \tag{31.14}$$

$Q_{\alpha\beta}$ est une matrice symétrique, puisque l'ordre des dérivations incluses dans (31.8) – voir aussi (31.6) – est sans importance. Par ailleurs, tout réseau de Bravais a un centre d'inversion[6] : $Q_{\alpha\beta}(\vec{R})$ est donc invariant dans le changement $\vec{R} \to -\vec{R}$; au total :

$$Q_{\alpha\beta}(\vec{R}) = Q_{\beta\alpha}(\vec{R}) = Q_{\alpha\beta}(-\vec{R}) \;; \tag{31.15}$$

compte tenu de ces symétries, les équations du mouvement (31.14) s'écrivent finalement :

$$\boxed{M\ddot{u}_\alpha(\vec{R},\, t) = -\sum_{\vec{R}'}\sum_{\beta=x,\,y,\,z} Q_{\alpha\beta}(\vec{R}-\vec{R}')\, u_\beta(\vec{R}',\, t)} \tag{31.16}$$

[6] Il en va donc de même pour tout cristal monoatomique formé d'objets ponctuels ou à symétrie sphérique (scalaires, par opposition à vecteurs, ...)

La résolution effective de ces équations suppose données des conditions initiales. Si le réseau contient N ($\gg 1$) atomes, il faut en principe connaître les $3N$ positions initiales, $\vec{u}(\vec{R},\,0)$ et les vitesses $\dot{\vec{u}}(\vec{R},0)$.

Tout ce qui précède est certes d'inspiration classique, mais le passage à la formulation quantique est immédiat, puisque les équations du mouvement de Heisenberg sont, pour un oscillateur harmonique, formellement exactement les mêmes que leurs équivalents classiques (ce calcul sera explicité pour la chaîne linéaire, voir section 31.2) ; cette identité formelle subsiste pour un ensemble d'oscillateurs en interaction, à la condition expresse que le couplage soit *bilinéaire* par rapport aux écarts (coordonnées), ce qui est le cas ici. Une fois tous les oscillateurs quantifiés, et compte tenu de :

$$\big[u_\alpha(\vec{R}),\, P_{\vec{r}(\vec{R}')\beta}\big] = \mathrm{i}\hbar\delta_{\alpha\beta}\delta_{\vec{R}\vec{R}'} \ , \tag{31.17}$$

on obtient :

$$M\ddot{u}_{\mathrm{H}\alpha}(\vec{R},\,t) = -\sum_{\vec{R}'}\ \sum_{\beta=x,\,y,\,z} Q_{\alpha\beta}(\vec{R}-\vec{R}')\,u_{\mathrm{H}\beta}(\vec{R}',\,t) \ , \tag{31.18}$$

où $\vec{u}_{\mathrm{H}}(\vec{R},t)$ est la représentation de Heisenberg de l'observable position du cœur vibrant près du site $\vec{R}$. Les conditions initiales sont maintenant la donnée d'un état de départ $|\Psi(0)\rangle$ pour le réseau, avec lequel on peut en principe calculer les valeurs moyennes $\langle\Psi(0)|\vec{u}_{\mathrm{H}}(\vec{R},0)|\Psi(0)\rangle$ et $\langle\Psi(0)|\dot{\vec{u}}_{\mathrm{H}}(\vec{R},0)|\Psi(0)\rangle$ ($\dot{\vec{u}}(\vec{R}) = \vec{P}_{\vec{u}(\vec{R})}/M$).

31.2 Modes normaux d'un réseau unidimensionnel

Afin d'illustrer les idées précédentes, considérons un réseau unidimensionnel de paramètre a, dont les atomes peuvent vibrer seulement le long du réseau. Un tel modèle est certes académique, mais des systèmes réels s'en rapprochent, quand les couplages harmoniques sont très anisotropes, provoquant une rigidité du réseau qui rendent négligeables les déplacements dans deux des trois directions de l'espace.

En vertu du caractère unidimensionnel, chaque déplacement $\vec{u}$ n'a qu'une composante et la matrice $K_{\alpha\beta}$ se réduit à un scalaire. La position de chaque ion est[7] r_j :

$$r_j = R_j + u_j \ , \qquad R_j = ja \qquad (j\in\mathbb{Z}) \ . \tag{31.19}$$

et (31.5) devient :

$$U_{\mathrm{h}} = \frac{1}{2}\sum_{j,\,j'\in\mathbb{Z}} \frac{1}{2!}\,K(R_j - R_{j'})(u_j - u_{j'})^2 \ . \tag{31.20}$$

Même dans ce cas très simple, le calcul effectif de la fonction $K(R)$ à partir des premiers principes est en général très difficile, pour ne pas dire impossible. L'usage est de s'en

[7] r_j n'est pas une longueur, c'est une abscisse $\in\mathbb{R}$.

remettre à une nouvelle approximation *ad hoc*, consistant à ne considérer que des couplages entre premiers voisins, ce qui n'introduit ici qu'un seul paramètre $K(\pm a)$, noté[8] simplement K :

$$K(R_j - R_{j'}) = \begin{cases} K & \text{si } |j-j'| = 1 \\ 0 & \text{autrement} \end{cases} \quad . \tag{31.21}$$

Il n'existe pas vraiment de justification théorique d'une telle approximation, ce d'autant plus que les interactions élastiques sont à longue portée, surtout en basse dimensionnalité. En fait, la validation de ce schéma minimal – à $D = 1$ comme en dimension supérieure – repose sur ses succès quand on le confronte aux résultats expérimentaux. Ceci étant admis, et comme les deux termes $j' = j \pm 1$ sont en fait égaux entre eux (ce qui fait sauter un facteur $\frac{1}{2}$), le potentiel harmonique de la chaîne linéaire se réduit à :

$$\boxed{U_\mathrm{h} = \frac{1}{2} K \sum_{j\in\mathbb{Z}} (u_j - u_{j+1})^2} \tag{31.22}$$

L'interprétation de cette expression est élémentaire ; les équations de Lagrange classiques avec une telle énergie potentielle conduisent à :

$$M\ddot{u}_j = -\frac{\partial U_\mathrm{h}}{\partial u_j} = -K(2u_j - u_{j-1} - u_{j+1}) \tag{31.23}$$

soit :

$$\boxed{\ddot{u}_j = -\omega_0^2 (2u_j - u_{j-1} - u_{j+1})} \tag{31.24}$$

où on a posé $K = M\omega_0^2$ et $\omega_0 = \sqrt{\frac{K}{M}} > 0$.

D'un autre point de vue, la force de rappel pour l'ion situé en ja résulte exclusivement des deux forces exercées par ses deux premiers voisins situés en $(j \pm 1)a$; elle vaut $-K(u_j - u_{j-1}) - K(u_j - u_{j+1})$: c'est très exactement ce qui figure au second membre de (31.24). Autrement dit, dans l'approximation en cours, tout se passe comme si on avait une chaîne de boules identiques reliées par des ressorts parfaits ayant tous la même constante de raideur K : on est finalement réduit à trouver les modes propres (classiques) d'un ensemble d'oscillateurs couplés, au nombre de N s'il y a N ions[9].

Si l'on regarde les choses d'un point de vue quantique, l'interprétation est la même ; le Hamiltonien de la chaîne d'ions est :

$$H = \sum_{j\in\mathbb{Z}} \frac{P_j^2}{2M} + \frac{1}{2} K \sum_{j\in\mathbb{Z}} (u_j - u_{j+1})^2 \ , \tag{31.25}$$

où P_j est le moment conjugué (impulsion) de l'ion situé au voisinage du point ja. Compte tenu de $[u_i,\, P_j] = \mathrm{i}\hbar\delta_{ij}$, les équations de Heisenberg sont :

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t} u_{j\mathrm{H}} = [u_j,\, H]_\mathrm{H} = \sum_{j\in\mathbb{Z}} \frac{1}{2M} [u_j,\, P_j^2]_\mathrm{H} = \frac{\mathrm{i}\hbar}{M} P_{j\mathrm{H}} \quad \Longleftrightarrow \quad \dot{u}_{j,\mathrm{H}} = \frac{P_{j\mathrm{H}}}{M} \ , \tag{31.26}$$

[8] À ne pas confondre avec un vecteur de $\tilde{\mathcal{B}}$!

[9] Le même problème a été rencontré dans le chapitre 25, en prémices de la quantification du champ électromagnétique (voir éq. (25.2)). Une occasion de plus de constater l'universalité de l'oscillateur harmonique...

et :

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}P_{j\mathrm{H}} = [P_j,\, H]_\mathrm{H} = \frac{1}{2}K\sum_{j'\in\mathbb{Z}}\left[P_j,\, (u_{j'} - u_{j'+1})^2\right]_\mathrm{H} \ . \qquad (31.27)$$

Le commutateur est nul si $j' \neq j,\, j-1$; la somme donne donc les deux seuls termes non-nuls :

$$\left[P_j,\, (u_j - u_{j+1})^2 + (u_{j-1} - u_j)^2\right]_\mathrm{H} = -2i\hbar(2u_j - u_{j+1} - u_{j-1}) \ , \qquad (31.28)$$

d'où $\frac{\mathrm{d}}{\mathrm{d}t}P_{j\mathrm{H}} = -K(2u_j - u_{j+1} - u_{j-1})$; avec (31.26), on trouve bien formellement la même équation que dans le cas classique :

$$\boxed{\ddot{u}_{j\mathrm{H}} = -\,\omega_0^2\,(2u_j - u_{j+1} - u_{j-1})_\mathrm{H}} \qquad (31.29)$$

qui est la version quantique de (31.24). La seule distinction entre ces deux équations est la nature des objets qui y figurent : des coordonnées $u_j(t)$ au sens classique, ou des observables $u_{j\mathrm{H}}(t)$ en représentation de Heisenberg. Dans un cas comme dans l'autre, on est en face d'un système linéaire de N inconnues, qui se traite toujours de la même façon, puisque, en définitive, tous ces objets commutent toujours entre eux : l'algèbre reste donc classique (commutative).

Finalement, notant x_j soit u_j, soit $u_{j\mathrm{H}}$, il convient de résoudre le système linéaire homogène :

$$\boxed{\ddot{x}_j = -\,\omega_0^2\,(2x_j - x_{j+1} - x_{j-1})} \qquad (31.30)$$

où, à ce stade, la nature précise des inconnues x_j est techniquement sans importance. Pour éviter des difficultés inessentielles, on procède comme d'habitude en considérant dans un premier temps un réseau de taille finie (N atomes) muni des conditions aux limites commodes précisées en temps utile. Dès lors, (31.30) est un système linéaire $N \times N$ du second ordre en temps, dont la solution physique exige la connaissance de $2N$ conditions initiales. Une méthode commode de résolution consiste à effectuer une transformée de Fourier en temps, en posant :

$$x_j(t) \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} X_j(\omega)\,\mathrm{e}^{-i\omega t}\,\mathrm{d}\omega \ ; \qquad (31.31)$$

en vertu de l'hermiticité – ou de la réalité – de x_j, on doit avoir :

$$X_j(-\omega) = X_j(\omega)^\dagger \qquad \text{ou} \qquad X_j(-\omega) = X_j(\omega)^* \ . \qquad (31.32)$$

Le report dans (31.30) fournit :

$$\int_{-\infty}^{+\infty} \left[\omega^2 X_j - \omega_0^2\,(2X_j - X_{j+1} - X_{j-1})\right]\mathrm{e}^{-i\omega t}\,\mathrm{d}\omega = 0 \quad \forall\omega \ , \qquad (31.33)$$

ce qui impose aux X_j de satisfaire :

$$\boxed{\omega^2 X_j - \omega_0^2\,(2X_j - X_{j+1} - X_{j-1}) = 0} \ . \qquad (31.34)$$

Il s'agit d'un système linéaire homogène, qui n'a de solution non triviale que si son déterminant est nul. Cette condition fixe une fois pour toutes les seules valeurs possibles de ω^2, en tant que zéros d'un polynôme de degré N, $\mathcal{P}_N(\omega^2)$. À ce stade – même si le calcul explicite n'est pas encore fait –, les valeurs possibles de ω^2 sont formellement déterminées[10] par l'équation $\mathcal{P}_N(\omega^2) = 0$.

D'une autre façon, le système homogène (31.34) est une équation aux différences finies, dont les solutions sont de la forme $X_j =$ (quelque chose)^j – et on peut toujours dire que (quelque chose) est une exponentielle. Posons donc :

$$X_j = C\mathrm{e}^{\mathrm{i}j\phi} \; ; \tag{31.35}$$

ϕ est une quantité réelle, sinon on aurait d'un côté ($j > 0$) ou de l'autre ($j < 0$) des divergences exponentielles inacceptables physiquement, tandis que C est un facteur indéterminé[11]. Alors (31.34) devient :

$$[\omega^2 - \omega_0^2\,(2 - \mathrm{e}^{\mathrm{i}\phi} - \mathrm{e}^{-\mathrm{i}\phi})]\,\mathrm{e}^{\mathrm{i}j\phi} = 0 \quad\Longleftrightarrow\quad \omega^2 - 4\omega_0^2\,\sin^2\frac{\phi}{2} = 0 \; . \tag{31.36}$$

ω étant formellement déjà déterminé, ceci fixe ϕ complètement, ou *vice versa.* En pratique[12], il est plus commode d'obtenir d'abord ϕ, en choisissant des conditions aux limites agréables dont la nature précise est sans importance pour les résultats physiques tant que $N \gg 1$; comme d'habitude, on prend des conditions cycliques identifiant X_{N+1} et X_1 :

$$X_{N+1} = X_1 \quad\Longleftrightarrow\quad \mathrm{e}^{\mathrm{i}(N+1)\phi} = \mathrm{e}^{\mathrm{i}\phi} \quad\Longleftrightarrow\quad \mathrm{e}^{\mathrm{i}N\phi} = 1 \; . \tag{31.37}$$

Comme déjà vu à plusieurs reprises, ceci quantifie les valeurs de ϕ, un angle défini à 2π près ; ϕ est sans dimension, X_j est relatif au site situé en ja : il est donc naturel d'introduire un vecteur d'onde $\vec{k}$, ici à une dimension[13], k, tel que $\phi = ka$. Alors :

$$\mathrm{e}^{\mathrm{i}Nka} = 1 \quad\Longleftrightarrow\quad k = \text{entier} \times \frac{2\pi}{Na} \quad\Longleftrightarrow\quad k = n\,\frac{2\pi}{Na} \equiv k_n \; . \tag{31.38}$$

Puisqu'il y a N modes propres au total, n prend N valeurs, par exemple[14] $1, 2, \ldots, N$; les valeurs de ω sont alors explicites :

$$\boxed{\omega = \pm\,\omega(k) \;, \quad \omega(k) = 2\omega_0 \left|\, \sin\frac{ka}{2} \,\right|} \tag{31.39}$$

[10] Il resterait à prouver que ce polynôme a non seulement tous ses zéros réels mais encore qu'ils sont tous positifs ou nuls. La réalité est assurée par le fait que $\mathcal{P}_N$ est associé à une matrice symétrique (hermitique réelle). La positivité peut se démontrer soit par récurrence en N, soit en utilisant des propriétés générales (pour une chaîne ouverte – non refermée *à la* Born - von Kármán –, $\mathcal{P}_N$ est un polynôme bien connu, appelé polynôme de Gegenbauer).

[11] Le système (31.34) étant *homogène*, les X_j ne sont fixés qu'à un facteur multiplicatif près.

[12] Les deux procédures sont strictement équivalentes pour la raison suivante. Appliquer les conditions cycliques, c'est finalement jouer implicitement avec les opérateurs de rotation d'angle entier $\frac{2\pi}{N}$, qui représentent des symétries de la chaîne refermée par l'identification $N + 1 = 1$. Ces opérateurs forment un groupe abélien et, ce faisant, on diagonalise complètement *ipso facto* la matrice associée au système linéaire (31.34).

[13] k est donc défini à $\frac{2\pi}{a}$ près.

[14] On pourrait tout autant prendre $n = -\frac{N}{2}, -\frac{N}{2}+1, \ldots, +\frac{N}{2}$, en supposant N pair – hypothèse non-restrictive dans la limite $N \gg 1$.

Au total, il n'y a bien que N valeurs distinctes pour k, ce qui donne également N valeurs distinctes pour les deux branches $\pm\,\omega(k)$, eu égard à la parité de $\omega(k)$. L'usage ordinaire[15] consiste à faire varier k symétriquement autour de l'origine, essentiellement entre $\pm\frac{N}{2}\frac{2\pi}{Na} = \pm\frac{\pi}{a}$, qui n'est autre que la première zone de Brillouin du réseau (voir fig. 31.1). La variation de la fonction $\omega(k)$ avec k constitue la loi (relation) de dispersion pour les modes de vibration ; pour un k donné, les X_j sont donnés par :

$$\boxed{X_j(\omega(k)) = C\mathrm{e}^{\mathrm{i}kja}} \tag{31.40}$$

La fréquence (donc l'énergie) de ces modes de vibration s'annule en $k = 0$, tout comme les magnons rencontrés au chapitre 30, section 30.6. Pour les mêmes raisons, ils joueront un rôle de tout premier plan pour la thermodynamique basse température : on verra (section 31.6) qu'ils sont les principaux responsables de l'écart à la loi classique de Dulong et Petit (voir aussi Tome I, pp. 111 et 139).

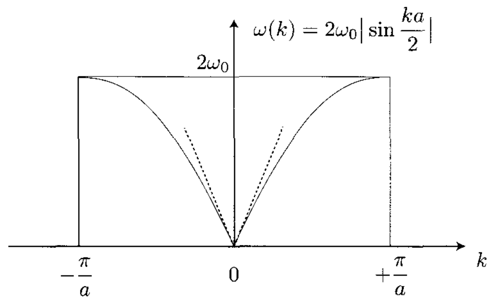


Figure 31.1: Relation de dispersion pour un réseau unidimensionnel avec un couplage entre premiers voisins. La pente des lignes en pointillés est la vitesse $c \stackrel{\text{déf}}{=} a\omega_0$.

Il est instructif d'examiner en détail la collection des déplacements u_j résultant d'un état initial précis. Il suffit pour cela d'inverser la relation de Fourier (31.31) ; les valeurs possibles de ω sont en fait discrètes[16], de sorte que l'intégrale dans (31.31) est à ce stade une somme $\sum_k$. Compte tenu des deux branches $\pm\omega(k)$ de la loi de dispersion, on peut toujours écrire $x_j(t)$ sous la forme :

$$x_j(t) = \sum_k X_j\big(\omega(k)\big)\left(\alpha_{k+}\mathrm{e}^{\mathrm{i}\omega(k)t} + \alpha_{k-}\mathrm{e}^{-\mathrm{i}\omega(k)t}\right) \ , \tag{31.41}$$

où les $\alpha_{k\pm}$ sont deux constantes d'intégration, à k fixé ; au total, il faut trouver $2N$ constantes $C\alpha_{k\pm}$, ce qui est possible en utilisant les $2N$ conditions initiales[17]. Compte tenu de (31.40), il vient :

$$x_j(t) = \sum_k C\mathrm{e}^{\mathrm{i}kja}\left(\alpha_{k+}\mathrm{e}^{\mathrm{i}\omega(k)t} + \alpha_{k-}\mathrm{e}^{-\mathrm{i}\omega(k)t}\right) \ . \tag{31.42}$$

[15] $\omega(k)$ est de toute façon une fonction périodique de k, avec la (plus petite) période $\frac{2\pi}{a}$.

[16] Ce sont les N solutions de $\mathcal{P}_N(\omega^2) = 0$.

[17] qui, dans le cas quantique, contiennent les valeurs moyennes dans l'état initial prescrit des opérateurs au sens de Schrödinger.

Dans le cadre classique, les x_j sont les u_j et selon (31.42) :

$$u_j(t) = \sum_k C\mathrm{e}^{\mathrm{i}kja}\big(\alpha_{k\,+}\mathrm{e}^{\mathrm{i}\omega(k)t} + \alpha_{k\,-}\mathrm{e}^{-\mathrm{i}\omega(k)t}\big) \ , \tag{31.43}$$

et les moments sont $(P_j \equiv M\dot{r}_j = M\dot{u}_j)$:

$$P_j(t) = M\sum_k \mathrm{i}\omega(k)\, C\mathrm{e}^{\mathrm{i}kja}\, \big(\alpha_{k\,+}\mathrm{e}^{\mathrm{i}\omega(k)t} - \alpha_{k\,-}\mathrm{e}^{-\mathrm{i}\omega(k)t}\big) \ . \tag{31.44}$$

Les $2N$ conditions initiales s'écrivent :

$$u_j(0) = \sum_{k\in[-\frac{\pi}{a},+\frac{\pi}{a}]} C\mathrm{e}^{\mathrm{i}kja}(\alpha_{k\,+} + \alpha_{k\,-}) \ , \tag{31.45}$$

$$P_j(0) = M \sum_{k\in[-\frac{\pi}{a},+\frac{\pi}{a}]} \mathrm{i}\omega(k)\, C\mathrm{e}^{\mathrm{i}kja}(\alpha_{k\,+} - \alpha_{k\,-}) \ . \tag{31.46}$$

Pour simplifier, supposons toutes les vitesses initiales nulles ; l'égalité (31.46) dit alors que $\alpha_{k\,+} = \alpha_{k\,-} = \alpha_k$ ce qui donne pour (31.43) :

$$u_j(t) = \sum_{k\in[-\frac{\pi}{a},+\frac{\pi}{a}]} 2\alpha_k C\mathrm{e}^{\mathrm{i}kja} \cos\omega(k)t \ , \tag{31.47}$$

avec maintenant :

$$u_j(0) = \sum_{k\in[-\frac{\pi}{a},+\frac{\pi}{a}]} 2\alpha_k C\mathrm{e}^{\mathrm{i}kja} \ . \tag{31.48}$$

Cette dernière relation s'inverse facilement en utilisant[18] $\sum_{j=1}^{N} \mathrm{e}^{\mathrm{i}(k-k')ja} = N\delta_{kk'}$, d'où :

$$2N\alpha_k C = \sum_{j=1}^{N} u_j(0)\, \mathrm{e}^{-\mathrm{i}kja} \ . \tag{31.49}$$

Ainsi, pour une chaîne classique abandonnée hors de son point d'équilibre sans vitesse initiale pour les atomes, les déplacements à un instant ultérieur sont donnés par :

$$u_j(t) = \frac{1}{N}\sum_{j'=1}^{N} u_{j'}(0) \sum_{k\in[-\frac{\pi}{a},+\frac{\pi}{a}]} \mathrm{e}^{\mathrm{i}k(j-j')a} \cos\omega(k)t \ ; \tag{31.50}$$

avec $N \gg 1$, la somme sur k est équivalente à l'intégrale : $\sum_{|k|\leq\pi/a} \rightarrow \frac{L}{2\pi}\int_{-\pi/a}^{+\pi/a} \mathrm{d}k$ et, comme $L = Na$:

$$\boxed{u_j(t) = \sum_{j'=1}^{N} u_{j'}(0)\frac{a}{2\pi}\int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}k(j-j')a} \cos\omega(k)t} \tag{31.51}$$

[18] Cette égalité s'obtient en calculant la somme de la progression géométrique, $k - k'$ étant de la forme entier $\times\ \frac{2\pi}{N}$.

Compte tenu de la dépendance de ω en k (voir (31.39)), l'intégrale porte sur des exponentielles du genre $\mathrm{e}^{\mathrm{i}k(j-j')a}\,\mathrm{e}^{\pm 2\mathrm{i}\omega_0 t\,\sin(ka/2)}$; le changement de variable $ka/2 = \phi$ fait apparaître des intégrales $\mathrm{e}^{2\mathrm{i}(j-j')\phi}\,\mathrm{e}^{\pm 2\mathrm{i}\omega_0 t\,\sin\phi}$, caractéristiques de la définition des fonctions de Bessel d'indice entier. Au total, $u_j(t)$ est la combinaison linéaire[19] :

$$\boxed{u_j(t) = \sum_{j'=1}^{N} u_{j'}(0)\,J_{2(j-j')}(2\omega_0 t)} \tag{31.52}$$

qui produit un champ de déplacements oscillant en temps et dans l'espace. En particulier, si seul l'atome en j_0 est déplacé au départ de u_0, $u_j(0) = u_0\delta_{jj_0}$, d'où :

$$u_j(t) = u_0\,J_{2(j-j_0)}(2\omega_0 t) \ ; \tag{31.53}$$

connaissant le comportement asymptotique des fonctions de Bessel aux grands temps[20] $t \gg \omega_0^{-1}$, on a à peu près :

$$\omega_0 t \gg 1 \ : \ u_j(t) \sim \frac{u_0}{\sqrt{\pi\omega_0 t}}\cos\left[2\omega_0 t - (j-j_0)\frac{\pi}{2} - \frac{\pi}{4}\right] \ ; \tag{31.54}$$

à t fixé et grand, le champ des déplacements a en gros l'allure d'une sinusoïde amortie, avec une amplitude déclinant $\sim t^{-\frac{1}{2}}$. Clairement, le paquet d'ondes *localisé* de la déformation initiale s'élargit et se dilue le long de la chaîne ; notamment, le déplacement initial en j_0 s'évanouit et tend vers zéro, mais très lentement. La même dynamique de relaxation a été rencontrée à propos d'une chaîne de spins, voir (30.217).

En $t = 0$, le second membre de (31.50) est[21] :

$$\frac{1}{N}\sum_{j'} u_{j'}(0)\sum_{k}\mathrm{e}^{\mathrm{i}k(j-j')a} = \frac{1}{N}\sum_{j'} u_{j'}(0)\,N\delta_{jj'} = u_j(0) \ , \tag{31.55}$$

comme il se doit. Par ailleurs, si on déplace initialement chaque atome de la même quantité $u(0)$, (31.50) donne :

$$u_j(t) = \frac{\dot{u}(0)}{N}\sum_{j'=1}^{N}\sum_{k}\mathrm{e}^{\mathrm{i}k(j-j')a}\,\cos\omega(k)t = \frac{u(0)}{N}\sum_{k}\mathrm{e}^{\mathrm{i}kja}\,N\delta_{ka\,0}\,\cos\omega(k)t = u(0) \ . \tag{31.56}$$

Dans ces circonstances, chaque atome garde sa position initiale à tout temps, ce qui est bien normal : un tel état de départ est en fait une translation en bloc du réseau, sans déformation, n'induisant aucune force de rappel et ne provoquant donc aucun mouvement ultérieur. Ceci permet de retrouver le fait que le mode $k = 0$ est en réalité associé à un déplacement (translation) en bloc, d'énergie forcément nulle.

[19]On se souvient que seule la fonction J_0 est non-nulle à l'origine (voir par exemple le développement (19.76)).

[20]Pour $x \gg 1$ (et $x \lesssim |n|$), on a $J_n(x) \sim \sqrt{\frac{2}{\pi x}}\cos(x - n\frac{\pi}{2} - \frac{\pi}{4})$.

[21]Il en va de même si on fait $t = 0$ au second membre de (31.51).

D'une façon générale, les modes de petit k ($ka \ll 1$) correspondent à des déformations dynamiques $u_j(t)$ variant lentement en fonction de l'entier j. Autrement dit, le champ donnant les déplacements des atomes a une variation spatiale *lente* à l'échelle du pas du réseau (si les déplacements initiaux $u_{j'}(0)$ possèdent cette propriété, il en ira évidemment de même à tout temps). Dès lors, ce sont les modes de grande longueur d'onde qui sont prépondérants : on ne peut bien sûr se contenter strictement de $k = 0$, mais, à l'inverse, le *voisinage* de $k = 0$ suffit pour la description à grande échelle. Alors, dans (31.51), on peut remplacer $\omega(k)$ par son développement limité (voir (31.39)) :

$$\boxed{\omega(k) = 2\omega_0\,\frac{|k|a}{2} + \ldots \equiv |k|c + \ldots \qquad c \stackrel{\text{déf}}{=} a\omega_0} \tag{31.57}$$

où a été introduite une certaine vitesse c, qui n'est autre que la vitesse du son (vitesse de groupe et vitesse de phase coïncident dans la limite du solide continu[22]). Cette quantité est la seule vitesse disponible avec les paramètres physiques du problème[23].

Posant $x = ja$, $u(x = ja,\, t) \stackrel{\text{déf}}{=} u_j(t)$, (31.51) devient :

$$u(x,t) \simeq \frac{a}{2\pi}\int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} \mathrm{d}k \sum_{j'=1}^{N} u(j'a,\, 0)\,\mathrm{e}^{\mathrm{i}k(x-j'a)}\,\cos kct\ . \tag{31.58}$$

Dans cette optique à grande échelle, tout se passe comme si le pas a du réseau était un "infiniment petit" ; dans la continuité de la description en cours, il est licite de remplacer $a\sum_{j'}$ par $\int \mathrm{d}x'$, l'expression (31.58) ayant essentiellement le même sens que[24] :

$$u(x,t) \simeq \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} \frac{\mathrm{d}k}{2\pi}\int_{\mathbb{R}} \mathrm{d}x'\,u(x',\, 0)\,\mathrm{e}^{\mathrm{i}k(x-x')}\,\cos kct\ . \tag{31.59}$$

L'une ou l'autre de ces dernières expressions montre que la déformation initiale se propage le long du réseau sous la forme d'un paquet d'ondes dispersé suivant la loi (31.39) linéarisée, soit (31.57) dont vitesse de phase et vitesse de groupe coïncident (limite des grandes longueurs d'onde). Une branche de vibration qui tend vers zéro linéairement en k est appelée *branche acoustique* :

$$\boxed{\lim_{k\to 0}\frac{\omega(k)}{k} = \mathrm{C^{ste}} \quad \Longleftrightarrow \quad \text{branche acoustique}} \tag{31.60}$$

[22]Dans la limite $ka \to 0$, la nature discrète du réseau s'efface, toutes les longueurs d'onde du champ des déplacements étant très grandes devant le pas a du réseau.

[23]Par comparaison avec une relation de dispersion du genre $\varepsilon(\vec{k}\,) = (p^2c^2 + m^2c^4)^{1/2}$, c'est-à-dire $\omega(\vec{k}\,) = c(k^2 + k_0^2)^{1/2}$ avec $k_0 \neq 0$, on dit que des excitations $\varepsilon(\vec{k}\,) \equiv \hbar\omega(\vec{k}\,) \propto |k|$ n'ont pas de *masse* (elles sont *massless*).

[24]La linéarisation en k de la loi de dispersion – et la limite $a \to 0$ – rend élémentaire l'intégration en k de (31.59), qui fait apparaître $\delta(x - x' \pm ct)$, d'où $u(x,t) = \frac{1}{2}\,[u(x - ct,\, t = 0) + u(x + ct,\, t = 0)]$; cette expression est bien la solution de l'équation de propagation $(c^{-2}\partial_{t^2} - \partial_{x^2})u = 0$ et représente la dynamique d'une onde dans un milieu non dispersif (toutes les fréquences avancent à la même vitesse) ; dans ces conditions, tout paquet d'ondes se propage en bloc, sans déformation. Retomber sur l'équation de propagation ne doit pas surprendre : le deuxième membre de (31.30) n'est rien d'autre qu'une dérivée seconde *discrète* par rapport à l'espace, un fait qui a été utilisé dans le chapitre 25 quand il s'est agi de donner l'exemple le plus simple de la quantification d'un champ classique (sous-section 25.1.1).

D'un autre côté, si les déplacements initiaux varient vite avec j, il faut conserver la loi de dispersion (31.39) dans toute son intégrité. La vitesse de groupe associée à la loi de dispersion (31.39) est :

$$v_{\mathrm{g}}(k) = \frac{\mathrm{d}\omega(k)}{\mathrm{d}k} = a\omega_0 \cos\frac{ka}{2} \ ; \tag{31.61}$$

elle s'annule lorsque $ka = \pi$, c'est-à-dire en bord de zone. Un paquet d'ondes ayant une amplitude maximum en $k = k_0 = \frac{\pi}{a}$ ne se propage donc pas et fait du sur-place. Pour comprendre ce que ceci signifie en termes de déplacements, choisissons la distribution initiale où :

$$u_j(t=0) = (-1)^j\, u_0 \ , \tag{31.62}$$

état où les sites pairs sont tous écartés d'une même quantité u_0, tous les sites impairs étant écartés de $-u_0$ (chaîne *dimérisée*) ; c'est bien un cas où le champ des écarts initiaux varie vite à l'échelle du réseau (il est nul en moyenne sur deux mailles adjacentes). Le report de (31.62) dans (31.51) fait apparaître la somme :

$$\sum_{j'=1}^{N} \mathrm{e}^{-\mathrm{i}j'ka}\, u_{j'}(0) = u_0 \sum_{j'=1}^{N} \mathrm{e}^{-\mathrm{i}j'ka}\,(-1)^j = u_0 \sum_{j'=1}^{N} \mathrm{e}^{\mathrm{i}j'(\pi-ka)} = u_0\, N\, \delta_{ka\,\pi} \ . \tag{31.63}$$

Insérée dans une intégrale sur k, la quantité $N\delta_{ka\,\pi}$ se comporte comme une fonction de Dirac, $C\,\delta(ka-\pi)$ – soit $\frac{C}{a}\,\delta(k-\frac{\pi}{a})$; en outre, il est facile de voir[25] que $C = 2\pi$. L'expression de $u_j(t)$ selon (31.51) est alors :

$$u_j(t) = u_0 \int_{-\frac{\pi}{a}}^{+\frac{\pi}{a}} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}jka}\, \delta(k-\frac{\pi}{a})\, \cos\omega(k)t = u_0\, \mathrm{e}^{\mathrm{i}j\pi}\, \cos 2\omega_0 t \ , \tag{31.64}$$

où l'expression (31.39) a été utilisée. En définitive, pour l'état initial (31.62), les déplacements à l'instant t sont donnés par :

$$\boxed{u_j(t) = u_j(0)\, \cos 2\omega_0 t} \tag{31.65}$$

Tous les sites vibrent donc à la même fréquence $2\omega_0$ et deux sites voisins restent en opposition de phase, tout comme ils l'étaient à $t = 0$. Si l'on regarde le paquet d'ondes réduit à deux tels sites voisins, on voit en effet qu'il reste en place : il ne bouge pas et ne se déforme pas. Cette *stationnarité* est compréhensible, eu égard au fait que cette déformation initiale n'attaque qu'un mode normal du réseau de dimères ; l'absence de non-linéarités ne peut briser la symétrie locale présente dans l'état de départ.

Le cas quantique se traite de la même façon ; en prenant les valeurs moyennes des opérateurs *à la* Heisenberg sur un état initial des atomes où toutes les vitesses sont nulles, on obtient exactement les mêmes expressions, $u_j(t)$ donné par (31.51) étant alors la valeur moyenne (quantique) du déplacement à l'instant t quand les noyaux sont immobiles (en moyenne) dans l'état initial.

[25] Le pas en k est $\delta k = \frac{2\pi}{a}$ et $L = Na$.

31.3 Modes normaux d'un réseau unidimensionnel à deux atomes par maille

On a vu dans la section 31.2 que les modes $k \approx 0$ d'une chaîne monoatomique correspondent à un champ de déplacement de grande longueur d'onde. La limite $k = 0$ donne un mode de translation en bloc d'énergie nulle : il n'est donc pas surprenant que la pulsation $\omega(k)$ s'annule en $k = 0$.

D'un autre point de vue, rien n'interdit de considérer le réseau précédent comme un réseau de maille $a' = 2a$, avec une base à deux atomes (identiques) séparés de la même distance a, donc constituée de molécules diatomiques (dimères) ; dans cette optique, le plus petit vecteur du réseau réciproque est de norme $\frac{2\pi}{2a}$, la première zone de Brillouin s'étendant entre $\pm\frac{\pi}{2a}$. En suivant le même procédé que pour les électrons dans un potentiel périodique, on peut récrire les résultats de la section 31.2 dans un schéma en zones réduites, avec $|k| \leq \frac{\pi}{2a}$ (voir fig. 31.2). La première branche qui apparaît, notée désormais ω_{ac} (acoustique) est donc identique à (31.39), mais est définie entre $\pm\frac{\pi}{2a}$:

$$\omega_{\text{ac}}(k) = 2\omega_0 \left| \sin\frac{ka}{2} \right| , \qquad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a} \right] . \tag{31.66}$$

Les arcs de (31.39) situés entre $\pm\frac{\pi}{2a}$ et $\pm\frac{\pi}{a}$ doivent être décalés vers la gauche et vers la droite d'un multiple de $\frac{2\pi}{2a}$, de façon à se retrouver entre $\pm\frac{\pi}{2a}$; les translatant de $\pm\frac{\pi}{a}$, on obtient une nouvelle branche $\omega_{\text{opt}}(k)$, appelée *branche optique* :

$$\omega_{\text{opt}}(k) \stackrel{\text{déf}}{=} 2\omega_0 \left| \sin\frac{(k \mp \frac{\pi}{a})a}{2} \right| \equiv 2\omega_0 \cos\frac{ka}{2} , \qquad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a} \right] . \tag{31.67}$$

Dans le schéma en zones réduites, la pulsation est maintenant *bivaluée*, les deux branches se rejoignant en $\pm\frac{\pi}{2a}$ – notamment, il y a deux modes $k = 0$. Celui de la branche acoustique représente toujours la translation en bloc et a naturellement une énergie nulle ; en revanche, celui de la branche optique correspond à une vibration en opposition de phase des deux membres du dimère, tous les dimères vibrant en phase (voir (31.65)) ; son énergie n'est pas nulle, évidemment. Un mode optique se différencie nettement d'un mode acoustique par le fait que sa pulsation ne tend pas vers zéro aux grandes longueurs d'onde puisqu'il représente, dans le schéma en zones réduites, le mode où la vitesse de groupe s'annule et où deux atomes voisins vibrent en opposition de phase.

La description qui précède semble artificielle lorsque les deux partenaires du dimère sont identiques, mais elle permet une transition vers le cas où ils sont, de fait, différents, une situation qui correspond à un vrai cristal avec deux atomes par maille. Dans l'immédiat, on va se borner à différencier un atome sur deux du réseau de pas a par sa masse. Le traitement peut être étendu à une base diatomique où les partenaires sont à une distance d l'un de l'autre (en pratique, $d < a$), voir problème 31.9.2.

Un raisonnement dans un cas limite permet de prévoir qualitativement les résultats. Les atomes n'ayant pas la même masse – il peut s'agir de deux isotopes (masses M_1

et M_2, $M_1 < M_2$) –, le potentiel harmonique (qui ne dépend que de la distribution électronique périphérique puisqu'il est obtenu dans l'approximation des cœurs fixes) est le même pour tous les sites, entraînant que la constante de raideur K est la même partout. En revanche, il y a maintenant deux pulsations :

$$\omega_0 = \sqrt{\frac{K}{M_1}} \ , \qquad \Omega_0 = \sqrt{\frac{K}{M_2}} < \omega_0 \ , \tag{31.68}$$

de sorte que, en présence des deux masses, *deux* branches sont à attendre.

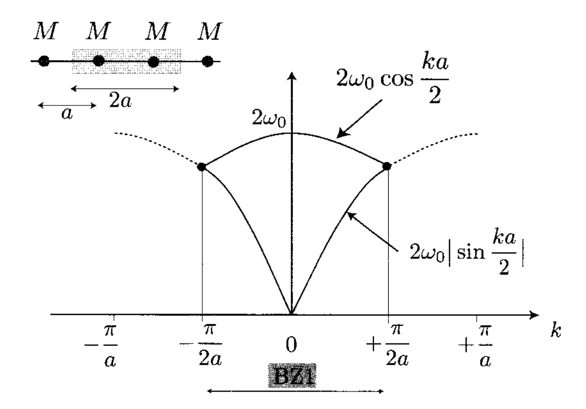


Figure 31.2: Relation de dispersion (31.39) dans le schéma en zones réduites pour le réseau de maille double avec deux atomes *identiques par maille.*

L'allure des lois de dispersion est facilement prévisible si les masses sont très différentes. Avec $M_1 \ll M_2$, chaque dimère a une masse à peu près égale à M_2, et le réseau de dimères a un pas égal à $2a$. En considérant chacun d'entre eux comme une grosse particule, on voit qu'il existe des modes de type acoustique, la loi étant approximativement[26] donnée par :

$$\omega_{\mathrm{ac}}(k) \simeq 2\frac{\Omega_0}{\sqrt{2}} \left|\sin\frac{k(2a)}{2}\right| \ , \qquad k \in \left[-\frac{\pi}{2a}, +\frac{\pi}{2a}\right] \qquad (M_1 \ll M_2) \ ; \tag{31.69}$$

c'est bien la pulsation $\frac{\Omega_0}{\sqrt{2}}$ qui doit être insérée dans (31.67), puisque chaque grosse particule a une masse à peu près égale à M_2 et que chacune des deux grandes masses est reliée à la petite par un ressort de constante K ; le ressort effectif[27] entre les deux atomes les plus massifs a une constante $K_{\text{eff}}^{-1} = K^{-1} + K^{-1} = 2K^{-1}$, soit $K_{\text{eff}} = K/2$. La fréquence d'oscillation est donc $\Omega_{\text{eff}}^2 = (K/2)M_2$, soit $\Omega_{\text{eff}} = \Omega_0/\sqrt{2}$. La branche part de 0 en $k = 0$ et atteint sa valeur maximale $\sqrt{2}\Omega_0$ en $k = \frac{\pi}{2a}$.

[26]On va évidemment trouver des corrections, puisque le rapport M_2/M_1 n'est pas infini. Noter que $M_1 \ll M_2$ est équivalent à $\Omega_0 \ll \omega_0$.

[27]Quand on met des ressorts en série, ce sont les inverses des constantes de raideur qui s'additionnent (au total, c'est plus *mou*) ; au contraire, si les ressorts sont mis en parallèle, c'est plus *dur* et les constantes de raideur s'additionnent. Une chaîne 1-d infinie a une compressibilité infinie (elle est infiniment molle, sa susceptibilité est infiniment grande).

Ceci définit une première classe de modes de vibration. Il existe aussi un mode de vibration intradimère, représentant cette fois une oscillation rapide de la petite masse M_1 rappelée par la grande masse M_2 quasiment immobile. En considérant un seul dimère, cette oscillation se fait strictement à la pulsation $\sqrt{2}\,\omega_0 \gg \Omega_0$, et a donc une grande énergie[28] ; en raison du couplage mécanique entre les petites masses[29], l'énergie correspondante dépendra de k, mais on s'attend à ce qu'elle soit peu dispersée autour de ω_0 – elle est donc relativement plate – et, en tout cas est un mode optique, sa pulsation tendant vers une valeur finie en $k = 0$.

En définitive, dans le cas de deux masses très différentes, $M_1 \ll M_2$, on attend deux branches, l'une optique (étroite), l'autre acoustique, qui ne se rejoignent plus en $k = \frac{\pi}{2a}$: un *gap* (relativement grand) s'est ouvert en bord de zone par comparaison au cas $M_1 = M_2$ traité comme ci-dessus (deux atomes identiques par maille). Pour deux masses très voisines, $M_1 \lesssim M_2$, des arguments analogues montrent que les dispersions des deux branches sont comparables, et que le *gap* est cette fois tout petit. Dans tous les cas, le mécanisme d'ouverture de *gaps* est exactement le même que pour les électrons dans un potentiel périodique, le cas $M_1 \lesssim M_2$ correspondant aux électrons presque libres.

Le calcul exact détaillé est facile, avec M_1 et M_2 quelconques (mais toujours $M_1 < M_2$ pour fixer les idées). On repart des équations du mouvement (31.24) adaptées au cas analysé maintenant ; les petites masses M_1 étant situées sur les sites impairs, les grandes masses M_2 sur les sites pairs, les équations du mouvement sont :

$$M_1\,\ddot{u}_{2j+1} = -\frac{\partial U_\mathrm{h}}{\partial u_{2j+1}} = -K\,(2u_{2j+1} - u_{2j} - u_{2j+2}) \ , \qquad (31.70)$$

et :

$$M_2\,\ddot{u}_{2j} = -\frac{\partial U_\mathrm{h}}{\partial u_{2j}} = -K\,(2u_{2j} - u_{2j-1} - u_{2j+1}) \ . \qquad (31.71)$$

Avec la même convention que précédemment (x_j est soit un déplacement classique $u(t)$, soit la représentation de Heisenberg $u_\mathrm{H}(t)$), on obtient le double système (c'est la généralisation du système unique (31.30)) :

$$\ddot{x}_{2j+1} = -\omega_0^2(2x_{2j+1} - x_{2j} - x_{2j+2}) \ , \quad \ddot{x}_{2j} = -\Omega_0^2(2x_{2j} - x_{2j-1} - x_{2j+1}) \ , \qquad (31.72)$$

que l'analyse de Fourier transforme en :

$$\omega^2 X_{2j+1} - \omega_0^2(2X_{2j+1} - X_{2j} - X_{2j+2}) = 0 \ , \ \omega^2 X_{2j} - \Omega_0^2(2X_{2j} - X_{2j-1} - X_{2j+1}) = 0 \ . \qquad (31.73)$$

[28]Pour une masse M_1 liée de part et d'autre à deux masses inertes ($u_{j\pm1} = 0$), l'équation du mouvement est $M_1\ddot{u}_j = -2Ku_j$, d'où le facteur $\sqrt{2}$.

Par ailleurs, la relation entre haute fréquence et haute énergie tient tant dans le cas classique que dans le cas quantique. Dans le cas quantique, elle résulte de $E = \hbar\omega$; dans le cas classique, elle vient du fait que l'énergie mécanique d'un oscillateur écarté de x_0 de sa position initiale sans vitesse est égale à $\frac{1}{2}m\omega_0^2x_0^2$, d'autant plus grande que ω_0 est grand. Pour un écart initial donné, l'énergie à dépenser est d'autant plus grande que le ressort est plus ferme.

[29]Comme les grosses masses M_2 ne sont pas totalement inertes, il existe un couplage résiduel indirect entre les petites masses M_1.

Il s'agit à nouveau d'équations aux différences finies dont les solutions sont les puissances d'exponentielles complexes. Les X sont des amplitudes de Fourier des déplacements ; comme les déplacements sont d'autant plus faibles que l'inertie est grande, et qu'il y a deux masses différentes, il convient d'introduire deux constantes de proportionnalité C_1 et C_2 pour les sites impairs et pairs respectivement. On pose donc :

$$X_{2j+1} = C_1 \mathrm{e}^{\mathrm{i}(2j+1)ka} \ , \qquad X_{2j} = C_2 \mathrm{e}^{\mathrm{i}2jka} \ . \tag{31.74}$$

Le report dans (31.73) donne le système 2×2 :

$$(\omega^2 - 2\omega_0^2)\, C_1 + 2\omega_0^2 \cos ka\, C_2 = 0 \ , \quad 2\Omega_0^2 \cos ka\, C_1 + (\omega^2 - 2\,\Omega_0^2)\, C_2 = 0 \ , \tag{31.75}$$

qui a une solution non-triviale si et seulement si le déterminant est nul ; ceci conduit à une équation bicarrée dont les solutions en ω^2 sont :

$$\omega^2 = \omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 + \Omega_0^2)^2 - 4\omega_0^2\Omega_0^2 \sin^2 ka} \ . \tag{31.76}$$

La combinaison sous la racine carrée est aussi $(\omega_0^2 - \Omega_0^2)^2 + 4\omega_0^2\Omega_0^2 \cos^2 ka$: elle est bien positive ; par ailleurs, la quantité ω^2 est visiblement positive quel que soit le signe au second membre. L'égalité (31.76) donne notamment les deux branches positives :

$$\boxed{\omega_\pm(k) \stackrel{\text{déf}}{=} \left[\omega_0^2 + \Omega_0^2 \pm \sqrt{(\omega_0^2 - \Omega_0^2)^2 + 4\omega_0^2\Omega_0^2 \cos^2 ka}\right]^{1/2}} \tag{31.77}$$

l'ensemble des relations de dispersion étant formé des quatre branches $\pm\omega_\pm(k)$ (signes indépendants). En $k = 0$, les $\omega_\pm(k)$ valent 0, $\sqrt{2(\omega_0^2 + \Omega_0^2)}$; plus précisément, un développement limité donne ($ka \sim 0$) :

$$\omega_-(k) \simeq \sqrt{2\frac{\omega_0^2\Omega_0^2}{\omega_0^2 + \Omega_0^2}}\, |k|a \ , \quad \omega_+(k) \simeq \sqrt{2(\omega_0^2 + \Omega_0^2)}\Big[1 - \frac{\omega_0^2\Omega_0^2}{2(\omega_0^2 + \Omega_0^2)^2}\,(ka)^2\Big] \ ; \tag{31.78}$$

ω_- est la branche acoustique (elle s'annule linéairement en k), ω_+ est la branche optique, parabolique près de $k = 0$, avec une concavité négative.

Près du bord de zone[30], $k \lesssim \frac{\pi}{2a}$, le développement limité de (31.77) donne :

$$\omega_+(k) \simeq \sqrt{2}\,\omega_0\Big[1 + \frac{\Omega_0^2}{2(\omega_0^2 - \Omega_0^2)}\big(\frac{\pi}{2} - ka\big)^2\Big] \ , \tag{31.79}$$

et :

$$\omega_-(k) \simeq \sqrt{2}\,\Omega_0\Big[1 - \frac{\omega_0^2}{2(\omega_0^2 - \Omega_0^2)}\big(\frac{\pi}{2} - ka\big)^2\Big] \ . \tag{31.80}$$

Les deux branches sont donc à peu près paraboliques en bord de zone, avec des concavités opposées, et sont séparées par le *gap* $\sqrt{2}\,(\omega_0 - \Omega_0)$, qui s'annule bien si on en revient à deux masses identiques – et alors la variation est *linéaire* dans le voisinage de $\frac{\pi}{2a}$. À

[30]Plus précisément, il faut $\frac{\omega_0}{\sqrt{\omega_0^2 - \Omega_0^2}}|ka - \frac{\pi}{2}| \ll 1$.

l'inverse, si $M_1 \ll M_2$ (soit $\omega_0 \gg \Omega_0$), (31.77) montre que la branche optique est peu dispersée ; en effet, on trouve alors :

$$\omega_+(k) \simeq \sqrt{2\,\omega_0^2\left(1+\frac{\Omega_0^2}{\omega_0^2}\cos^2 ka\right)} \ , \tag{31.81}$$

et comme $\frac{\Omega_0}{\omega_0} \ll 1$, la variation induite par $\cos^2 ka$ est faible : la bande optique est presque plate. Tous ces résultats quantitatifs confirment l'analyse intuitive préalable.

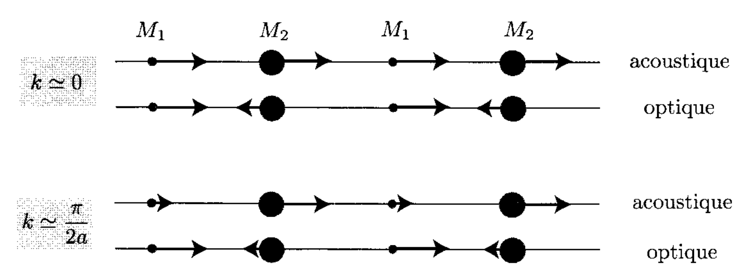


Figure 31.3: Représentation schématique des déplacements à un instant donné pour un réseau à deux atomes par maille.

Une fois obtenues les valeurs propres, il est possible d'analyser les mouvements respectifs des différents atomes dans chacun des modes obtenus (voir fig. 31.3). Il suffit pour cela d'examiner le rapport $\frac{C_1}{C_2}$, donné (par exemple) par :

$$\left(\frac{C_1}{C_2}\right)_\pm = -\frac{2\omega_0^2\cos ka}{\omega_\pm^2 - 2\omega_0^2} = \frac{2\cos ka}{1-\mu\mp\sqrt{(1+\mu)^2 - 4\mu\sin^2 ka}} \ , \qquad \mu \stackrel{\text{déf}}{=} \frac{\Omega_0^2}{\omega_0^2} = \frac{M_1}{M_2} \leq 1 \ . \tag{31.82}$$

En utilisant les développements limités de $\omega_\pm(k)$, il est facile d'obtenir les comportements du rapport en centre et en bord de zone pour chacun des deux modes :

- $k \to 0$:

mode optique : $$\left(\frac{C_1}{C_2}\right)_+ \simeq -\frac{1}{\mu}\left[1 - \frac{1-\mu}{2(1+\mu)}\,(ka)^2\right] = -\frac{M_2}{M_1}\left[1+\mathcal{O}(k^2a^2)\right] \ . \tag{31.83}$$

Les deux atomes vibrent en opposition de phase – ce que l'on attend d'un mode optique, avec une amplitude inversement proportionnelle à leur masse ; ce rapport reflète (à l'envers) les inerties : $C_1 > C_2$ si $M_1 < M_2$. Pour le mode acoustique ω_-, on trouve tout naturellement que les deux amplitudes sont du même ordre de grandeur par rapport au rapport des masses :

mode acoustique : $$\left(\frac{C_1}{C_2}\right)_- \simeq 1 - \frac{M_2 - M_1}{2(M_2+M_1)}\,(ka)^2 \ . \tag{31.84}$$

- $k \to \frac{\pi}{2a}$:
  $\omega_+^2 - 2\omega_0^2$ s'annule quadratiquement par rapport à $k - \frac{\pi}{2a} < 0$. Au contraire, le numérateur de (31.82) s'annule linérairement. Il en résulte que le rapport $(\frac{C_1}{C_2})_+$ diverge ; très précisément, un développement limité donne :

$$\text{mode optique :} \qquad \left(\frac{C_1}{C_2}\right)_+ \simeq \left(\frac{M_2}{M_1} - 1\right)\left(k - \frac{\pi}{2a}\right)^{-1} \longrightarrow -\infty\ . \qquad (31.85)$$

Dans le mode optique, près du bord de zone, l'atome lourd est presque immobile et seul l'atome léger vibre avec une amplitude notable (comme le montre le signe $-$, le mouvement est toujours en opposition de phase). C'est un mode de haute énergie, correspondant pratiquement à un mode de vibration intramoléculaire.

Pour ω_- (mode acoustique), il en va tout autrement car le dénominateur a maintenant une limite finie :

$$\text{mode acoustique :} \qquad \left(\frac{C_1}{C_2}\right)_- \simeq \frac{M_2}{M_2 - M_1}\left(\frac{\pi}{2} - ka\right) > 0\ ; \qquad (31.86)$$

cette fois, c'est l'atome massif qui a une grande amplitude, cependant que, relativement, l'atome léger est à peu près fixe ; l'énergie de vibration est nettement plus petite que pour le mode optique.

La figure 31.4 résume ces résultats.

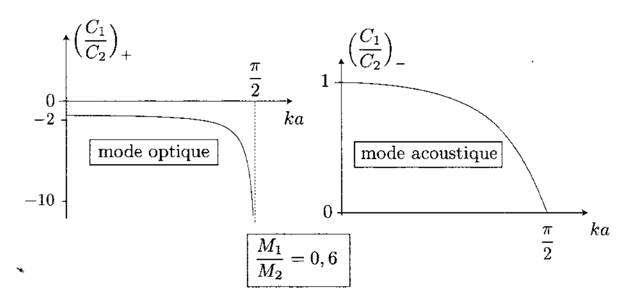


Figure 31.4: Rapport des amplitudes de vibration $\frac{C_1}{C_2}$ des deux masses M_1 et M_2 (voir (31.82)) ; les courbes sont tracées pour $\mu \stackrel{\text{déf}}{=} \frac{M_1}{M_2} = 0,6$.

31.4 Modes normaux d'un réseau tridimensionnel

Les vibrations d'un réseau tridimensionnel s'obtiennent exactement comme précédemment, en tenant compte simplement du fait que le déplacement de chaque atome est déterminé par un vecteur $\vec{u}(\vec{R})$.

31.4.1 Réseau sans base

Lorsque le réseau est sans base, les composantes de $\vec{u}(\vec{R})$ obéissent aux équations du mouvement (31.16). Tout comme avant, des solutions particulières (complètes) peuvent s'écrire sous la forme d'ondes planes $\vec{f}(\vec{R},\, t)$, dont il faut maintenant de surcroît préciser la polarisation $\vec{\varepsilon}$:

$$\vec{f}(\vec{R},\, t) = \vec{\varepsilon}\,\mathrm{e}^{\mathrm{i}(\vec{k}.\vec{R}-\omega t)} \ ; \tag{31.87}$$

pour chaque composante ε_α de $\vec{\varepsilon}$, le report de cette expression dans (31.16) donne :

$$-M\omega^2\,\varepsilon_\alpha\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}} = -\sum_\beta\sum_{\vec{R}'} Q_{\alpha\beta}(\vec{R}-\vec{R}')\,\varepsilon_\beta\,\mathrm{e}^{\mathrm{i}\vec{k}.\vec{R}'} \ , \tag{31.88}$$

égalité qui se récrit comme suit :

$$M\omega^2\,\varepsilon_\alpha = \sum_\beta\sum_{\vec{R}''} Q_{\alpha\beta}(\vec{R}'')\varepsilon_\beta\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}''} \equiv \sum_\beta \mathcal{Q}_{\alpha\beta}(\vec{k})\,\varepsilon_\beta \ , \tag{31.89}$$

ou encore, sous forme matricielle :

$$\boxed{\left[M\omega^2\mathbf{1} - \mathcal{Q}(\vec{k})\right]\vec{\varepsilon} = 0} \tag{31.90}$$

$\mathcal{Q}(\vec{k})$ est appelé *matrice dynamique.* N désignant toujours le nombre d'atomes, il y a $3N$ modes normaux distincts[31]. En vertu des symétries de la matrice $Q(\vec{R})$ (voir (31.15)) :

$$\mathcal{Q}(\vec{k}) \equiv \frac{1}{2}\Big[\sum_{\vec{R}} Q(\vec{R})\,\mathrm{e}^{-\mathrm{i}\vec{k}.\vec{R}} + \sum_{\vec{R}} Q(-\vec{R})\,\mathrm{e}^{+\mathrm{i}\vec{k}.\vec{R}}\Big] = \sum_{\vec{R}} Q(\vec{R})\,\cos\vec{k}.\vec{R} \ , \tag{31.91}$$

soit, compte tenu de (31.12) :

$$\mathcal{Q}(\vec{k}) = \sum_{\vec{R}} Q(\vec{R})\,\cos\vec{k}.\vec{R} - \sum_{\vec{R}} Q(\vec{R}) = -\,2\sum_{\vec{R}} Q(\vec{R})\,\sin^2\frac{\vec{k}.\vec{R}}{2} \ . \tag{31.92}$$

$\mathcal{Q}(\vec{k})$ est donc une combinaison réelle de matrices $Q(\vec{R})$ symétriques réelles de dimension égale à 3, c'est donc une matrice ayant les mêmes caractéristiques. En tant que telle, elle possède trois vecteurs propres réels (orthogonaux) satisfaisant :

$$\mathcal{Q}(\vec{k})\,\vec{\varepsilon}_\lambda = q_\lambda(\vec{k})\,\vec{\varepsilon}_\lambda \qquad (\lambda = 1,\, 2,\, 3) \ ; \tag{31.93}$$

dans la suite, on supposera éventuellement que ces vecteurs sont de surcroît normalisés, $\vec{\varepsilon}_\lambda \,.\, \vec{\varepsilon}_{\lambda'} = \delta_{\lambda\lambda'}$. Il y a ainsi trois polarisations distinctes (et orthogonales) pour un même

[31]On admet que le cristal ne peut pas tourner – pas de degrés de liberté de rotation. Les trois modes de translation sont inclus dans la suite ($\vec{k} = \vec{0}$). De toute façon, avec $N \gg 1$, $3N$ ou $3N-6$ font le même effet.

vecteur[32] $\vec{k}$. Les pulsations propres $\omega_\lambda(\vec{k})$ sont données par :

$$M\,\omega_\lambda^2(\vec{k}) = q_\lambda(\vec{k}) \quad \Longleftrightarrow \quad \omega_\lambda(\vec{k}) = \sqrt{\frac{q_\lambda(\vec{k})}{M}} \ . \tag{31.94}$$

Tout comme pour le réseau monoatomique unidimensionnel, les trois branches s'annulent linéairement en $\vec{k} = 0$; dans cette limite :

$$\mathcal{Q}(\vec{k}) \simeq -\frac{1}{2}\sum_{\vec{R}} Q(\vec{R})\,(\vec{k}.\vec{R})^2 \equiv -\frac{1}{2}k^2 \sum_{\vec{R}} Q(\vec{R})\,(\vec{\kappa}.\vec{R})^2 \ , \qquad k \to 0 \tag{31.95}$$

où $\vec{\kappa}$ est le vecteur unitaire porté par $\vec{k}$. Conformément à (31.94), les pulsations sont données par la racine carrée des valeurs propres de la matrice $-\frac{k^2}{2M}\sum_{\vec{R}} Q(\vec{R})\,(\vec{\kappa}.\vec{R})^2$; dans la limite $k \to 0$, elles dépendent donc linéairement du module de $\vec{k}$, et de sa direction par l'intermédiaire de $\vec{\kappa}$. En définitive, on peut toujours les écrire :

$$\boxed{\omega_\lambda(\vec{k}) \simeq k\,c_\lambda(\vec{\kappa}) \qquad k \to 0} \tag{31.96}$$

égalité servant de définition à $c_\lambda(\vec{\kappa})$, vitesse (en général anisotrope) de propagation du son dans le cristal.

31.4.2 Réseau avec base

Lorsque le réseau contient une base (*molécule*) à n atomes, il existe au total $3nN$ modes de vibrations, c'est-à-dire $3n$ modes pour chaque valeur de $\vec{k}$. Parmi ces derniers, il y en a forcément 3 dont la fréquence tend vers zéro à petit $\vec{k}$, donnant à la limite la translation d'énergie nulle. Il reste donc $3(n-1)$ modes, qui sont de nature optique, leur fréquence ayant une valeur finie en $\vec{k} = 0$. Ils sont en fait issus des $3(n-1)$ modes de vibration-rotation d'une molécule, qui présentent une dispersion en $\vec{k}$ compte tenu de l'interaction entre les molécules constituant le cristal.

31.5 Quantification des vibrations de réseau

Une fois effectuée l'analyse classique des modes normaux de vibration du réseau, leur quantification est immédiate. Il suffit de revenir au Hamiltonien (31.2) où U est remplacé par la forme quadratique (31.8). Pour un réseau sans base, les $3N$ composantes $u_\alpha(\vec{R})$ peuvent être exprimées à l'aide des déplacements normaux $\xi_\lambda(\vec{k})$ qui diagonalisent la

[32] La disposition relative des polarisations $\vec{\varepsilon}_\lambda$ et de $\vec{k}$ dépend de la symétrie (ponctuelle) du cristal. Pour un cristal isotrope, pour chaque $\vec{k}$, on peut toujours choisir les solutions de sorte qu'il existe un mode longitudinal et deux modes tranverses. Lorsque le cristal est anisotrope, il n'en va plus ainsi, sauf si $\vec{k}$ est lui-même invariant dans certaines opérations de symétrie.

matrice dynamique (31.90). La transformation correspondante étant unitaire, la partie cinétique T de (31.2) est invariante en forme et devient :

$$T = \sum_{\lambda,\,\vec{k}} \frac{\Pi_{\lambda\vec{k}}^2}{2M} \; , \qquad \Pi_{\lambda\vec{k}} \stackrel{\text{déf}}{=} -\mathrm{i}\hbar\,\frac{\partial}{\partial\xi_{\lambda\vec{k}}} \; , \tag{31.97}$$

cependant que la partie potentielle est diagonale par construction :

$$U_{\mathrm{h}} = \sum_{\lambda,\,\vec{k}} \frac{1}{2}\, M\omega_\lambda^2(\vec{k})\,\xi_{\lambda\vec{k}}^2 \; . \tag{31.98}$$

Les degrés de liberté normaux $\xi_{\lambda\vec{k}}$ et leurs moments conjugués $\Pi_{\lambda\vec{k}}$ satisfont les relations de commutation canoniques :

$$[\xi_{\lambda\vec{k}},\, \Pi_{\lambda'\vec{k}'}] = \mathrm{i}\hbar\,\delta_{\lambda\lambda'}\,\delta_{\vec{k}\vec{k}'} \; , \tag{31.99}$$

en conséquence du fait que le passage des $\{u\}$ aux $\{\xi\}$ s'effectue par une transformation unitaire, qui ne change pas la forme des relations fondamentales de commutation. H est maintenant une *somme* de $\simeq 3N$ Hamiltoniens d'oscillateurs harmoniques :

$$H = \sum_{\lambda,\,\vec{k}} \left[\frac{\Pi_{\lambda\vec{k}}^2}{2M} + \frac{1}{2}\, M\omega_\lambda^2(\vec{k})\,\xi_{\lambda\vec{k}}^2\right] \; . \tag{31.100}$$

L'additivité de H signifie que chacun des oscillateurs est indépendant de tous les autres. De ce fait, la construction des états de vibration du cristal s'effectue en formant les produits des fonctions propres d'oscillateurs harmoniques $\psi_{\lambda\vec{k}}$, associées à l'énergie $\hbar\omega_\lambda(\vec{k})$ du mode. Chaque *quantum* d'excitation est appelé *phonon* ; si le mode $(\lambda,\, \vec{k})$ est au fondamental, il contient zéro phonon et a pour énergie $\hbar\omega_\lambda(\vec{k})/2$, s'il est dans le premier état excité, il y a un phonon de polarisation λ et d'énergie $\hbar\omega_\lambda(\vec{k})$, etc. : les phonons sont au champ des vibrations de réseau les analogues des photons pour le champ électromagnétique[33]. Notons de plus que, à petit $\vec{k}$, un phonon acoustique a une relation de dispersion *linéaire*, comme le photon : pour le coup, *mutatis mutandis*, l'analogie va jusqu'à l'identité formelle, qui se traduit par des comportements en température identiques tant que seuls les phonons acoustiques sont pertinents, soit à basse température – et tant que l'on peut ignorer la participation des électrons du solide aux excitations thermiques.

En définitive, l'état vibrationnel du cristal est complètement spécifié en donnant le nombre $n_{\lambda\vec{k}}$ de phonons dans chaque mode. Comme rien n'interdit un état de vibration arbitrairement excité pour un mode donné, $n_{\lambda\vec{k}}$ pouvant prendre n'importe quelle valeur dans $\mathbb{N}$, on peut caser un nombre arbitraire de phonons dans un seul mode : du point de vue de la statistique, les phonons sont des *bosons*[34]. Si le mode $(\lambda,\, \vec{k})$ contient $n_{\lambda\vec{k}}$

[33]Toutefois, le champ de phonons contient un mode *propagatif* longitudinal.

[34]Ceci se voit encore plus directement en introduisant les opérateurs d'annihilation $b_{\lambda\vec{k}}$ et $b_{\lambda\vec{k}}^\dagger$ de création pour chaque mode, équivalents des a et $a^\dagger$ de l'oscillateur harmonique élémentaire, dont on sait qu'ils satisfont $[a,\, a^\dagger] = \mathbf{1}$. Pour cette raison, on a ici :

$$[b_{\lambda\vec{k}},\, b_{\lambda'\vec{k}'}^\dagger] = \delta_{\lambda\lambda'}\,\delta_{\vec{k}\vec{k}'} \; , \tag{31.101}$$

relation de commutation caractéristique des bosons.

phonons, il a l'énergie $\hbar\omega_\lambda(\vec{k})(n_{\lambda\vec{k}} + \frac{1}{2})$. Pour une distribution des nombres d'occupation $n_{\lambda\vec{k}}$ sur l'ensemble des modes, l'énergie de vibration totale du réseau est :

$$E = \sum_{\lambda,\,\vec{k}} \hbar\omega_\lambda(\vec{k})\Big(n_{\lambda\vec{k}} + \frac{1}{2}\Big) \ ; \tag{31.102}$$

le demi-quantum est une simple constante additive, délaissée dans la suite puisque seules comptent les *différences* d'énergie. Dans l'approximation harmonique, l'expression retenue définitivement pour l'énergie totale de vibration est ainsi :

$$\boxed{E = \sum_{\lambda,\,\vec{k}} n_{\lambda\vec{k}}\, \hbar\omega_\lambda(\vec{k})} \tag{31.103}$$

Ayant obtenu le Hamiltonien, et ayant élucidé la nature des excitations, il est possible de décrire les propriétés statistiques ; la section suivante focalise sur le problème de la chaleur spécifique, dont le comportement est incompréhensible dans un cadre classique.

31.6 Contribution des phonons à la chaleur spécifique

Pour trouver la chaleur spécifique due aux phonons, il suffit d'écrire l'expression de la valeur moyenne de l'énergie du réseau en équilibre thermique à la température T. La distribution des nombres moyens d'occupation est celle de bosons, donnée par la statistique de Bose - Einstein :

$$n_{\lambda\vec{k}} = \frac{1}{\mathrm{e}^{\beta\hbar\omega_\lambda(\vec{k})} - 1} \ , \qquad \beta = \frac{1}{k_{\mathrm{B}}T} \ . \tag{31.104}$$

L'énergie moyenne (au sens de la statistique quantique) à la température T est alors donnée par (31.102) où l'expression (31.104) est utilisée. Si V désigne le volume de l'échantillon, la chaleur spécifique due aux phonons est[35] :

$$c_{V\,\mathrm{ph}} = \frac{1}{V}\frac{\partial E}{\partial T} = \frac{1}{V}\frac{\partial}{\partial T}\sum_{\lambda,\,\vec{k}} n_{\lambda\vec{k}}\,\hbar\omega_\lambda(\vec{k}) = \frac{1}{V}\frac{\partial}{\partial T}\sum_{\lambda,\,\vec{k}} \frac{\hbar\omega_\lambda(\vec{k})}{\mathrm{e}^{\beta\hbar\omega_\lambda(\vec{k})} - 1} \ . \tag{31.105}$$

À très haute température ($\beta\hbar\omega_\lambda(\vec{k}) \ll 1$) :

$$c_{V\,\mathrm{ph}} \simeq \frac{1}{V}\frac{\partial}{\partial T}\sum_{\lambda,\,\vec{k}} \hbar\omega_\lambda(\vec{k})\,(\beta\hbar\omega_\lambda(\vec{k}))^{-1} = \frac{1}{V}\frac{\partial}{\partial T}\sum_{\lambda,\,\vec{k}} k_{\mathrm{B}}T = \frac{3N}{V}k_{\mathrm{B}} \ . \tag{31.106}$$

C'est le résultat classique (loi de Dulong et Petit) : comme toujours, la limite haute température efface le *quantum* porté par $\hbar$ et la physique devient classique. Les corrections

[35]Sur cet exemple de calcul, on voit bien que la constante additive d'écart entre (31.102) et (31.103) ne joue aucun rôle.

Par ailleurs, comme il s'agit de chaleur *spécifique*, c_V est une grandeur intensive, et est homogène à $\mathrm{L}^{-D}k_{\mathrm{B}}$ pour un cristal D-dimensionnel.

quantiques sont faciles à écrire à partir de (31.105), mais sont souvent non-significatives car du même ordre de grandeur que les anharmonicités, ignorées d'un bout à l'autre.

L'autre limite, basse température, est évidemment plus intéressante puisque révélatrice des effets quantiques. Il est plus commode de discuter la forme intégrale de (31.105) ; pour un cristal de dimensionnalité D, et comme $\delta\vec{k} = (2\pi/L)^D$, il vient :

$$c_{V\,\mathrm{ph}} = \frac{\partial}{\partial T}\sum_\lambda \int_{\mathrm{BZ1}} \frac{\mathrm{d}^D k}{(2\pi)^D}\,\frac{\hbar\omega_\lambda(\vec{k})}{\mathrm{e}^{\beta\hbar\omega_\lambda(\vec{k})}-1} \ . \tag{31.107}$$

En se plaçant à température assez basse, on peut négliger tous les modes optiques éventuels[36], qui donnent des contributions exponentiellement petites par rapport à celle des modes acoustiques. Par ailleurs, en supposant de plus $k_\mathrm{B}T \ll \hbar\frac{\pi}{a}c_\lambda$ pour toutes les branches acoustiques retenues, on peut remplacer les vraies lois de dispersion $\hbar\omega_\lambda(\vec{k})$ par leurs expressions approchées à petit $\vec{k}$: à nouveau, la substitution n'introduit que des erreurs exponentiellement petites[37], ce qui autorise, dans la foulée, à rejeter à l'infini les bornes d'intégration. En conséquence, une expression admissible de $c_{V\,\mathrm{ph}}$ est :

$$c_{V\,\mathrm{ph}} = \frac{\partial}{\partial T}\sum_{\mathrm{acoustiques}} \int_{\mathbb{R}^D} \frac{\mathrm{d}^D k}{(2\pi)^D}\,\frac{\hbar k c_{\mathrm{ac}}(\vec{\kappa})}{\mathrm{e}^{\beta\hbar k c_{\mathrm{ac}}(\vec{\kappa})}-1} \ , \tag{31.108}$$

où la somme ne porte plus que sur les modes acoustiques, caractérisés à petit $\|\vec{k}\|$ par les différentes vitesses $c_{\mathrm{ac}}(\vec{\kappa})$, celles-ci ne dépendant que de la direction de $\vec{k}$, pas de son module. En séparant les dépendances radiale et angulaire, chaque intégrale de la somme se transforme comme suit ($\mathrm{d}^D k = k^{D-1}\mathrm{d}k\,\mathrm{d}\Omega_D$) :

$$\frac{\Omega_D}{(2\pi)^D}\int_0^{+\infty} k^{D-1}\mathrm{d}k \int_{\Sigma_D} \frac{\mathrm{d}\Omega(\vec{\kappa})}{\Omega_D}\,\frac{\hbar k c_{\mathrm{ac}}(\vec{\kappa})}{\mathrm{e}^{\beta\hbar k c_{\mathrm{ac}}(\vec{\kappa})}-1} \ , \tag{31.109}$$

où Σ_D est la sphère unité dans $\mathbb{R}^D$, et Ω_D sa surface[38]; l'intégrale sur k est toujours bien définie quel que soit $D>0$: l'intégrand est $\sim k^{D-1}$ en $k\sim 0$, et chute exponentiellement à l'infini ; aucune divergence (infrarouge ou ultraviolette) n'est donc à craindre.

Comme $c_{\mathrm{ac}}(\vec{\kappa})$ ne dépend que de l'orientation de $\vec{k}$, on peut extraire la fraction contenant la température de l'intégrale sur $\mathrm{d}\Omega$; posant $x = \beta\hbar k c_{\mathrm{ac}}(\vec{\kappa})$, il vient ainsi :

$$\frac{\Omega_D}{(2\pi)^D\beta^{D+1}\hbar^D}\int_0^{+\infty}\frac{x^D\mathrm{d}x}{\mathrm{e}^x-1}\int_{\Sigma_D}\frac{\mathrm{d}\Omega(\vec{\kappa})}{\Omega_D}\,\frac{1}{[c_{\mathrm{ac}}(\vec{\kappa})]^D} \ . \tag{31.111}$$

[36] présents s'il existe un motif.
Comme cela a été brièvement discuté à propos des magnons, les excitations *massives* (avec *gap*) deviennent non pertinentes pourvu que la température soit assez basse. En revanche, les excitations *massless* ne peuvent jamais être ignorées : c'est le cas des phonons acoustiques.

[37] L'exponentielle $\mathrm{e}^{\beta\hbar\omega_\lambda(\vec{k})}$ du dénominateur agit comme une fonction de coupure aux grands k et ne retient que les vecteurs d'onde tels que $\hbar\omega_\lambda(\vec{k}) \lesssim k_\mathrm{B}T$. Si $k_\mathrm{B}T \ll \hbar\frac{\pi}{a}c_\lambda$, la partie non-linéaire de la relation de dispersion est invisible.

[38] $\Omega_D = \frac{2\pi^{D/2}}{\Gamma(D/2)}$. Un peu plus loin dans ce calcul, on utilise la formule dite *de doublement* ([38], § 8.335.1) :

$$\Gamma(D) = \frac{2^{D-1}}{\sqrt{\pi}}\Gamma(\frac{D}{2})\Gamma(\frac{D+1}{2}) \ . \tag{31.110}$$

L'intégrale angulaire vaut ce qu'elle vaut, mais peut servir de base pour la définition de la moyenne angulaire de c_{ac}^{-D}, tous phonons (acoustiques) confondus ; on pose ainsi :

$$\frac{1}{c^D} \stackrel{\text{déf}}{=} \frac{1}{D} \sum_{\text{acoustiques}} \int_{\Sigma_D} \frac{\mathrm{d}\Omega(\vec{\kappa})}{\Omega_D} \frac{1}{[c_{\text{ac}}(\vec{\kappa})]^D} \ . \tag{31.112}$$

Dans (31.111), l'intégrale sur x se calcule en écrivant l'intégrand $x^D \mathrm{e}^{-x}(1-\mathrm{e}^{-x})^{-1}$, et en développant en série entière ; on trouve ainsi :

$$\int_0^{+\infty} \frac{x^D}{\mathrm{e}^x - 1}\,\mathrm{d}x = \Gamma(D+1)\sum_{n=1}^{+\infty} \frac{1}{n^{D+1}} \equiv \Gamma(D+1)\zeta(D+1) \ , \tag{31.113}$$

où Γ et ζ sont respectivement les fonctions d'Euler (de seconde espèce) et de Riemann. En utilisant (31.110), on trouve finalement :

$$\boxed{c_{V\,\text{ph}} = \frac{D^2(D+1)}{\pi^{(D+1)/2}}\Gamma(\frac{D+1}{2})\zeta(D+1)\left(\frac{k_{\text{B}}T}{\hbar c}\right)^D k_{\text{B}}} \tag{31.114}$$

Comme prévu, la chaleur spécifique ne présente aucune anomalie, même à basse dimensionnalité, puisque toutes les fonctions sont analytiques pour $\Re D > 0$. Pour un cristal tridimensionnel, $D = 3$, la chaleur spécifique basse température due aux phonons est donnée par l'expression[39] :

$$\boxed{c_{V\,\text{ph}} = \frac{2\pi^2}{5}\left(\frac{k_{\text{B}}T}{\hbar c}\right)^3 k_{\text{B}}} \tag{31.115}$$

Ainsi, pour un solide *ordinaire*, la contribution des phonons à la chaleur spécifique varie essentiellement comme T^3. Toutefois, ce comportement n'est observé qu'à *très* basse température : comme les lois de dispersion s'éloignent assez vite de la linéarité, il existe en général un grand intervalle de température où, la loi de Dulong et Petit étant déjà invalide, la loi en T^3 n'est pas pour autant vérifiée.

Pour un métal, cette variation typique en T^3 s'ajoute à la contribution du gaz de Fermi constitué par les électrons, qui donne un comportement *linéaire* en température, ce que l'on peut retrouver qualitativement comme suit : seuls les électrons situés près de la surface de la mer de Fermi sont susceptibles d'être excités thermiquement, et leur nombre est d'ordre $\mathcal{D}(\varepsilon_{\text{F}})k_{\text{B}}T$, $\mathcal{D}(\varepsilon_{\text{F}})$ étant la densité d'états au niveau de Fermi. La variation d'énergie du métal par fluctuation thermique est donc $\sim \mathcal{D}(\varepsilon_{\text{F}})k_{\text{B}}T \times k_{\text{B}}T$; par dérivation, on en déduit $c_{V\,\text{el}} \propto T$. Au total, la chaleur spécifique d'un métal à basse température est une loi du genre :

$$\boxed{c_V = (\gamma T + AT^3)\,k_{\text{B}}} \tag{31.116}$$

où γ est appelé *constante de Sommerfeld* (elle contient notamment la température de Fermi T_{F} du gaz d'électrons).

[39] $\Gamma(2) = 1$, $\zeta(4) = \frac{\pi^4}{90}$.

♦ *Remarques*

1. L'identité formelle relevée entre phonons *acoustiques* et *photons* (voir p. 1500) permet d'affirmer sans aucun calcul supplémentaire que l'énergie du rayonnement électromagnétique (un gaz parfait de photons) à l'équilibre, est proportionnelle à T^4 (dans $\mathbb{R}^3$). C'est l'essence de la loi de Stefan, un résultat également démontré par Boltzmann (voir Tome I, éq. (4.40)) par de purs arguments de Thermodynamique.

2. Comme dit dans le Tome I, pp. 111 et 139, l'un des succès de l'hypothèse de Planck est le calcul par Einstein de la chaleur spécifique d'un solide, dans l'hypothèse où les vibrations du solide ne mettent en jeu qu'un seul *quantum* de vibration (plus ou moins équivalent à un phonon *optique* (modèle d'Einstein). C'est pourquoi, si la tendance est bonne (c_V tend vers zéro à basse température contrairement à ce qu'affirme la loi classique de Dulong et Petit), le comportement est exponentiel, pour les raisons dites plus haut, et non pas en puissances de T, ce fait étant caractéristique d'une relation de dispersion tendant vers zéro avec k suivant une loi-puissance (voir problème 30.7.7). ♦

31.7 Manifestation des phonons dans les spectres de diffraction

La théorie de la diffraction exposée au chapitre 29, sous-section 29.3.2, repose sur l'hypothèse d'un cristal parfait (pas de défauts, ni structurels, ni chimiques), dont les objets physiques sont *immobiles*. La conclusion majeure alors obtenue, en ne considérant que la diffusion élastique, est l'apparition d'une tache ponctuelle de diffraction à chaque fois que le transfert de moment $\vec{q} \overset{\text{déf}}{=} \vec{k}_\text{f} - \vec{k}_\text{i}$ coïncide avec un vecteur $\vec{K}$ du réseau réciproque.

Cette conclusion doit évidemment être reconsidérée si les diffuseurs ont la liberté de se mouvoir, notamment s'ils effectuent des petites oscillations autour d'une position d'équilibre, définie par le réseau de Bravais. Cette nécessité s'impose quand on se souvient que c'est précisément l'ordre à longue distance, parfaitement déterminé pour un réseau sans défauts et statique, qui produit l'effet coopératif donnant lieu à des taches ponctuelles de diffraction. Par ailleurs, lorsqu'il s'agit de diffusion neutronique, il y a en outre possibilité de diffusion *inélastique* avec création ou annihilation de phonons, puisque les bons neutrons ont une énergie thermique et que, dès lors, on peut satisfaire les deux lois de conservation énergie et impulsion[40]. Afin d'en rester au niveau le plus simple, on continue toutefois pour l'instant à ne considérer que la diffusion élastique ; les principes et l'utilité de la diffusion inélastique sont brièvement discutés dans la section 31.8.

L'objet central est l'amplitude de diffusion $f(\vec{k}_\text{i}, \vec{k}_\text{f})$, dont on a vu qu'elle contenait

[40]En ce qui concerne l'irradiation X, on continue à laisser de côté les diffusions inélastiques à très haute énergies ($\gtrsim$ 20 keV).

le facteur décisif (voir chapitre 29, éq. (29.58)) :

$$S(\vec{q}) \stackrel{\text{déf}}{=} \sum_{\vec{R} \in \mathcal{B}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}} \, , \qquad \vec{q} = \vec{k}_\mathrm{f} - \vec{k}_\mathrm{i} \, . \tag{31.117}$$

La condition de von Laue (voir (29.65)) fixe un ensemble de directions de diffusion associées à des angles θ parfaitement déterminés, au sens où l'incertitude angulaire intrinsèque $\delta\theta$ est $\propto N^{-2}$ si N est le nombre de diffuseurs, nombre toujours macroscopique ($\sim 10^{23}$) en pratique.

Dans le cas où les atomes effectuent des petites vibrations, et avec les notations introduites en (31.1), on doit maintenant écrire :

$$S(\vec{q}) = \sum_{\vec{R} \in \mathcal{B}} \mathrm{e}^{-\mathrm{i}\vec{q}.[\vec{R}+\vec{u}(\vec{R})]} \, . \tag{31.118}$$

L'intensité de diffraction est proportionnelle à $|S(\vec{q})|^2$:

$$|S(\vec{q})|^2 = \sum_{\vec{R},\, \vec{R}'} \mathrm{e}^{\mathrm{i}\vec{q}.[\vec{R}+\vec{u}(\vec{R})]} \, \mathrm{e}^{-\mathrm{i}\vec{q}.[\vec{R}'+\vec{u}(\vec{R}')]} = \sum_{\vec{R},\, \vec{R}'} \mathrm{e}^{\mathrm{i}\vec{q}.(\vec{R}-\vec{R}')} \, \mathrm{e}^{\mathrm{i}\vec{q}.[\vec{u}(\vec{R})-\vec{u}(\vec{R}')]} \; ; \tag{31.119}$$

ceci donne l'intensité pour une configuration *donnée* des écarts $\vec{u}(\vec{R})$ mais ce que l'on observe, c'est bien évidemment la *moyenne*[41] thermique à une température donnée T. Ainsi, pour un réseau en équilibre thermique, la quantité importante à trouver est :

$$\langle |S(\vec{q})|^2 \rangle = \sum_{\vec{R},\, \vec{R}'} \mathrm{e}^{\mathrm{i}\vec{q}.(\vec{R}-\vec{R}')} \, \langle \mathrm{e}^{\mathrm{i}\vec{q}.[\vec{u}(\vec{R})-\vec{u}(\vec{R}')]} \rangle \, , \tag{31.120}$$

où les crochets représentent la moyenne prise au sens de la statistique (classique à très haute température, forcément quantique autrement). La moyenne de droite est invariante dans toute translation du réseau, de sorte que :

$$\boxed{\langle |S(\vec{q})|^2 \rangle = N \sum_{\vec{R}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}} \, \langle \mathrm{e}^{\mathrm{i}\vec{q}.\vec{u}(\vec{0})} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{u}(\vec{R})} \rangle} \tag{31.121}$$

Si tous les écarts étaient nuls, la somme serait proportionnelle à N, et l'intensité à N^2, comme on l'a noté dans la Remarque 1 p. 1354 ; pour trouver l'effet des phonons sur la diffraction de Bragg, il faut en tout premier lieu obtenir la quantité :

$$\boxed{\mathsf{Q} \stackrel{\text{déf}}{=} \langle \mathrm{e}^{\mathrm{i}\vec{q}.\vec{u}(\vec{0})} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{u}(\vec{R})} \rangle} \tag{31.122}$$

puis effectuer la sommation sur $\vec{R}$.

On sait que la coordonnée et le moment conjugué d'un oscillateur harmonique peuvent s'exprimer en combinaison linéaire des opérateurs de création et d'annihilation

[41] La fréquence typique d'une vibration est de l'ordre de 10^{12} Hz.

(voir Tome I, éq. (16.99)), ici notés $b^\dagger$ et b. Les déplacements $\vec{u}(\vec{R})$ – et leurs moments conjugués – sont plus précisément des sommes ou des différences des opérateurs $b^\dagger_{\vec{k}\lambda}$ et $b_{\vec{k}\lambda}$ associés aux modes normaux déterminés plus haut, de sorte que Q est de la forme :

$$\mathsf{Q} = \langle \mathrm{e}^{A}\mathrm{e}^{B}\rangle \ , \tag{31.123}$$

où $A \stackrel{\text{déf}}{=} \mathrm{i}\vec{q}.\vec{u}(\vec{0})$ et $B \stackrel{\text{déf}}{=} -\mathrm{i}\vec{q}.\vec{u}(\vec{R})$ sont des formes linéaires des $b^\dagger_{\vec{k}\lambda}$ et $b_{\vec{k}\lambda}$. Comme ces opérateurs satisfont $[b_{\vec{k}\lambda},\, b^\dagger_{\vec{k}'\lambda'}] = \delta_{\vec{k}\vec{k}'}\,\delta_{\lambda\lambda'}$, pour chaque facteur e^A et e^B, l'exponentielle de la somme est égale au produit des exponentielles. Q est ainsi égal à :

$$\mathsf{Q} = \langle \prod_{\vec{k},\,\lambda} \mathrm{e}^{A_{\vec{k}\lambda}}\mathrm{e}^{B_{\vec{k}\lambda}}\rangle \ . \tag{31.124}$$

Les deux opérateurs $A_{\vec{k}\lambda}$ et $B_{\vec{k}\lambda}$ étant des combinaisons linéaires des deux seuls opérateurs $b^\dagger_{\vec{k}\lambda}$ et $b_{\vec{k}\lambda}$, ils relèvent de la formule de Glauber (Tome I, problème 13.9.8, éq. (13.215)) :

$$\mathrm{e}^{A_{\vec{k}\lambda}+B_{\vec{k}\lambda}} = \mathrm{e}^{A_{\vec{k}\lambda}}\mathrm{e}^{B_{\vec{k}\lambda}}\mathrm{e}^{-\frac{1}{2}\,[A_{\vec{k}\lambda},\,B_{\vec{k}\lambda}]} \ , \tag{31.125}$$

où le commutateur est un simple nombre (qui, notamment, est insensible aux opérations de moyennes). Lue à l'envers, cette relation donne :

$$\mathrm{e}^{A_{\vec{k}\lambda}}\mathrm{e}^{B_{\vec{k}\lambda}} = \mathrm{e}^{A_{\vec{k}\lambda}+B_{\vec{k}\lambda}}\mathrm{e}^{+\frac{1}{2}\,[A_{\vec{k}\lambda},\,B_{\vec{k}\lambda}]} \ , \tag{31.126}$$

d'où :

$$\mathsf{Q} = \prod_{\vec{k},\,\lambda} \langle \mathrm{e}^{A_{\vec{k}\lambda}+B_{\vec{k}\lambda}}\rangle\, \mathrm{e}^{+\frac{1}{2}\,[A_{\vec{k}\lambda},\,B_{\vec{k}\lambda}]} \ . \tag{31.127}$$

$A_{\vec{k}\lambda} + B_{\vec{k}\lambda}$ est aussi une combinaison linéaire des deux opérateurs $b^\dagger_{\vec{k}\lambda}$ et $b_{\vec{k}\lambda}$:

$$A_{\vec{k}\lambda} + B_{\vec{k}\lambda} = \gamma_{\vec{k}\lambda}b_{\vec{k}\lambda} + \mu_{\vec{k}\lambda}b^\dagger_{\vec{k}\lambda} \ , \tag{31.128}$$

de sorte que l'on doit finalement calculer des valeurs moyennes du genre $\mathsf{q} \stackrel{\text{déf}}{=} \langle \mathrm{e}^{\gamma b+\mu b^\dagger}\rangle$ où b et $b^\dagger$ sont relatifs à un certain mode. Introduisant la fonction de partition canonique d'un oscillateur, Z (Tome I, éqs. (4.97) et (16.155)), on a :

$$\mathsf{q} = Z^{-1}\sum_{n=0}^{+\infty} \mathrm{e}^{-\beta\hbar\omega(n+\frac{1}{2})}\langle n|\mathrm{e}^{\gamma b+\mu b^\dagger}|n\rangle \ , \qquad Z = \sum_{n=0}^{+\infty} \mathrm{e}^{-\beta\hbar\omega(n+\frac{1}{2})} \ , \tag{31.129}$$

un calcul facile montre que[42] :

$$\boxed{\mathsf{q} \stackrel{\text{déf}}{=} \langle \mathrm{e}^{\gamma b+\mu b^\dagger}\rangle = \mathrm{e}^{\frac{1}{2}\langle(\gamma b+\mu b^\dagger)^2\rangle}} \tag{31.130}$$

[42] La relation (31.130) est connue sous le nom d'identité de Bloch, voir par exemple [235]. On peut y voir un avatar élaboré du fait que la transformée de Fourier d'une gaussienne est une gaussienne (voir exercice 31.9.5).

Alors ($j \equiv \vec{k}, \lambda$) :

$$\mathsf{Q} = \prod_j \mathrm{e}^{\frac{1}{2}\langle(\gamma_j b_j + \mu_j b_j^\dagger)^2\rangle}\,\mathrm{e}^{+\frac{1}{2}[A_j,\, B_j]} = \prod_j \mathrm{e}^{\frac{1}{2}\langle(A_j+B_j)^2\rangle}\,\mathrm{e}^{+\frac{1}{2}[A_j,\, B_j]} \ . \tag{31.131}$$

Maintenant :

$$\prod_j \mathrm{e}^{\frac{1}{2}\langle(A_j+B_j)^2\rangle} = \mathrm{e}^{\frac{1}{2}\sum_j\langle(A_j+B_j)^2\rangle} = \mathrm{e}^{\frac{1}{2}\langle[\sum_j A_j+B_j]^2\rangle} \ , \tag{31.132}$$

la dernière égalité venant du fait que, si $j \neq j'$, $\langle A_j B_{j'}\rangle = \langle A_j\rangle\langle B_{j'}\rangle$, et que la valeur moyenne d'une fonction linéaire de b et/ou $b^\dagger$ est nulle. (31.131) donne alors :

$$\mathsf{Q} = \mathrm{e}^{\frac{1}{2}\langle[\sum_j A_j+B_j]^2\rangle}\,\mathrm{e}^{+\frac{1}{2}\sum_j[A_j,\, B_j]} \ . \tag{31.133}$$

$[A_j,\, B_j]$ est un scalaire et peut être reporté dans l'argument de la première exponentielle, ce dernier devenant :

$$\frac{1}{2}\sum_j\langle(A_j+B_j)^2\rangle + [A_j,\, B_j] = \frac{1}{2}\sum_j\langle A_j^2 + B_j^2 + A_jB_j + B_jA_j + A_jB_j - B_jA_j\rangle =$$

$$\frac{1}{2}\sum_j\langle A_j^2 + B_j^2 + 2A_jB_j\rangle \tag{31.134}$$

Au total, la quantité Q définie en (31.123) est égale à :

$$\mathsf{Q} = \mathrm{e}^{\frac{1}{2}\sum_j\langle A_j^2+B_j^2+2A_jB_j\rangle} \ , \tag{31.135}$$

de sorte que $\mathsf{Q} \stackrel{\text{déf}}{=} \langle \mathrm{e}^{\mathrm{i}\vec{q}.\vec{u}(\vec{0})}\mathrm{e}^{-\mathrm{i}\vec{q}.\vec{u}(\vec{R})}\rangle$ est donné par :

$$\mathsf{Q} = \exp\Big\langle -\frac{1}{2}[\vec{q}.\vec{u}(\vec{0})]^2 - \frac{1}{2}[\vec{q}.\vec{u}(\vec{R})]^2 + [\vec{q}.\vec{u}(\vec{0})]\,[\vec{q}.\vec{u}(\vec{R})]\Big\rangle \ ; \tag{31.136}$$

les deux premiers termes de l'argument de l'exponentielle ont la même valeur moyenne, puisque tous les sites sont équivalents, $\langle[\vec{q}.\vec{u}(\vec{R})]^2\rangle = \langle[\vec{q}.\vec{u}(\vec{0})]^2\rangle$, d'où :

$$\mathsf{Q} = \mathrm{e}^{-2W}\,\exp\big\langle\,[\vec{q}.\vec{u}(\vec{0})]\,[\vec{q}.\vec{u}(\vec{R})]\big\rangle \ , \tag{31.137}$$

où on a introduit la quantité $2W$:

$$\boxed{2W \stackrel{\text{déf}}{=} \langle[\vec{q}.\vec{u}(\vec{0})]^2\rangle} \tag{31.138}$$

Traditionnellement, e^{-2W} s'appelle le *facteur de Debye - Waller*. Finalement, la quantité centrale (31.121) pilotant la diffusion élastique en présence des phonons a la forme :

$$\boxed{\langle|S(\vec{q})|^2\rangle = N\,\mathrm{e}^{-2W}\sum_{\vec{R}}\mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}}\,\mathrm{e}^{\Phi(\vec{R})}} \tag{31.139}$$

où la fonction $\Phi(\vec{R}) \stackrel{\text{déf}}{=} \big\langle[\vec{q}.\vec{u}(\vec{0})]\,[\vec{q}.\vec{u}(\vec{R})]\big\rangle$ est la fonction de corrélation spatiale (non centrée) des écarts d'équilibre (mécanique) projetés sur le transfert de moment.

Ce résultat montre les deux effets des vibrations de réseau sur les spectres de diffraction. D'une part, l'apparition du facteur de Debye - Waller révèle une *atténuation* des intensités de diffraction, d'autant plus grande que la moyenne $\langle[\vec{q}.\vec{u}(\vec{0})]^2\rangle$ est élevée, c'est-à-dire que la température est haute puisque l'écart quadratique moyen d'un oscillateur augmente avec la température (voir (31.143)) : par ce seul fait, les intensités des taches de Bragg décroissent quand la température augmente.

D'autre part, la somme de réseau figurant dans (31.139) ne conduit plus à une superposition *à poids égaux* de fonctions de Dirac, en vertu de la dépendance spatiale introduite par $\mathrm{e}^{\Phi(\vec{R})}$. Si cette fonction – qui est constante et égale à 1 pour le réseau immobile – vaut 1 pour la plupart des valeurs de $\vec{R}$, on retrouve en gros la situation du réseau fixe, les angles spécifiques de diffraction étant juste entachés d'un *flou* se traduisant par un petit élargissement angulaire des taches de diffraction. À l'inverse, si $\mathrm{e}^{\Phi(\vec{R})}$ varie significativement vite à l'échelle du réseau, l'incertitude angulaire $\delta\theta$ devient de l'ordre de 2π ; dit autrement, la diffusion coopérative perd tout caractère d'hypersélectivité[43]. Le point crucial est donc finalement l'allure de la fonction $\Phi(\vec{R})$.

À l'infini, cette fonction vaut 1 puisque pour deux points très éloignés l'un de l'autre, toute corrélation disparaît de sorte que la moyenne du produit est égal au produit des moyennes ; chaque moyenne valant 0, $\mathrm{e}^{\Phi(\infty)} = 1$. À l'opposé, en $\vec{R} = \vec{0}$, c'est la valeur quadratique moyenne des déplacements qui apparaît : $\mathrm{e}^{\Phi(\vec{0})} = \mathrm{e}^{\langle[\vec{q}.\vec{u}(\vec{0})]^2\rangle}$. La question se résume donc à savoir sur quelle distance la fonction $\Phi(\vec{R})$ passe d'une valeur à l'autre. On observe que ceci se produit presque toujours sur quelques distances atomiques à peine, ce qui signifie physiquement que la corrélation des écarts est à courte portée ; en première approximation, Φ est presque constante : pour l'immense majorité des valeurs de $\vec{R}$, $\mathrm{e}^{\Phi(\vec{R})} \simeq \mathrm{e}^{\Phi(\infty)}$. Cela étant, la somme dans (31.139) se décompose commme suit :

$$\sum_{\vec{R}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}}\,\mathrm{e}^{\Phi(\vec{R})} = \sum_{\vec{R}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}}\,\mathrm{e}^{\Phi(\infty)} - \sum_{\vec{R}} \mathrm{e}^{-\mathrm{i}\vec{q}.\vec{R}}\,[\mathrm{e}^{\Phi(\infty)} - \mathrm{e}^{\Phi(\vec{R})}] \ ; \tag{31.140}$$

la première somme est d'ordre N pour un réseau fini, et donne un produit de trois peignes de Dirac si $N \to +\infty$, reproduisant la condition de von Laue (ou de Bragg) ; au contraire, la deuxième somme est d'ordre quelques unités, et est en conséquence négligeable devant la première. Il en résulte que l'incertitude angulaire reste très petite et, dans les cas usuels, quasi-inobservable : en pareil cas, l'effet des vibrations se borne à une simple réduction d'intensité due au facteur de Debye - Waller, sans élargissement angulaire perceptible.

Cette conclusion doit cependant être tempérée eu égard à des effets de dimensionnalité : si elle est presque toujours correcte pour des cristaux à trois dimensions, elle peut être invalidée à $D = 2$ (ou *a fortiori* pour un réseau unidimensionnel), pour les raisons suivantes. La valeur $\Phi(\vec{0}) = \langle[\vec{q}.\vec{u}(\vec{0})]^2\rangle$ implique la moyenne du carré d'une quantité qui

[43] On raconte que von Laue, le pionnier de la diffraction X, ne croyait pas trop à la possibilité d'observer la diffusion sélective, justement en raison des effets de vibration de réseau. Si l'expérience s'est révélée positive – pour les cristaux tridimensionnels –, c'est *justement* parce que, en pratique, la fonction $\Phi(\vec{R})$ varie lentement.

est à nouveau une combinaison linéaire d'opérateurs de création et d'annihilation. Pour un seul oscillateur de fréquence ω, on a :

$$u = \sqrt{\frac{\hbar}{2M\omega}}\,(b + b^\dagger)\ . \tag{31.141}$$

La moyenne de u^2 est donc donnée par :

$$\langle u^2\rangle = \frac{\hbar}{2M\omega}\,\frac{1}{Z}\sum_{n=0}^{+\infty} \mathrm{e}^{-\beta[n+\frac{1}{2}]\hbar\omega}\,\langle b^2 + b^{\dagger 2} + bb^\dagger + b^\dagger b\rangle\ ; \tag{31.142}$$

un calcul simple donne le résultat[44] :

$$\boxed{\langle u^2\rangle = \frac{\hbar}{2M\omega}\coth\frac{\beta\hbar\omega}{2}} \tag{31.143}$$

Pour les combinaisons linéaires que constituent les $\vec{q}.\vec{u}(\vec{0})$, on va donc voir apparaître des arguments en exponentielle qui sont des sommes sur l'ensemble des modes de quantités analogues au second membre de (31.143) :

$$q^2\sum_{\vec{k}\lambda}\frac{\hbar}{2M\omega_\lambda(\vec{k})}\coth\frac{\beta\hbar\omega_\lambda(\vec{k})}{2}\ ; \tag{31.144}$$

pour chaque branche, la somme fait intervenir une intégrale du genre :

$$\int\frac{k^{D-1}\mathrm{d}k}{\omega_\lambda(\vec{k})}\coth\frac{\beta\hbar\omega_\lambda(\vec{k})}{2}\ , \tag{31.145}$$

où D est la dimension spatiale du réseau. Cette intégrale présente une singularité en $\vec{k} = 0$ en dimension inférieure à 3. En effet, près de l'origine, l'intégrand varie comme $k^{D-1}\omega_\lambda(\vec{k})^{-2} \sim k^{D-3}$; l'intégrale ne converge donc que si $D - 3 > -1$ soit $D > 2$. À $D = 2$, la divergence est *douce* (logarithmique), à $D = 1$, elle est relativement sévère. On sait traiter ces divergences *infra-rouges* en introduisant des coupures appropriées, ce qui revient à prendre en compte explicitement la *taille* de l'échantillon, forcément finie. Dans ces systèmes de basse dimensionnalité, il y a donc des effets de taille finie : les phonons y jouent un rôle important, donnant une véritable largeur angulaire aux taches de diffraction[45].

[44] À la limite classique, T finie, $\beta\hbar\omega \to 0$, on retrouve le théorème d'équipartition, $\frac{1}{2}M\omega^2\langle u^2\rangle = \frac{1}{2}k_\mathrm{B}T$, comme il se doit. À l'inverse, dans la limite ultra-quantique ($\hbar$ fixée, $T \to 0$), on obtient $\langle u^2\rangle = \frac{\hbar}{2M\omega}$: c'est l'écart quadratique de la position pour un oscillateur dans son état fondamental.

Encore un autre exemple de non-commutation des limites...

[45] Finalement, ce sont les phonons acoustiques qui donnent une amplification spectaculaire de l'effet vibrationnel. Ceci n'est pas surprenant : la diffusion sélective étant un phénomène coopératif impliquant l'ordre à longue distance, les phonons de grande longueur d'onde sont les plus capables de modifier sérieusement le comportement par rapport au réseau fixe.

31.8 Mesure des relations de dispersion des phonons : diffusion inélastique des neutrons

Un faisceau incident sur une cible est bien sûr *a priori* capable d'*exciter* la cible. En pareilles circonstances, les deux lois de conservation s'écrivent :

$$\hbar\vec{k}_\mathrm{i} + \vec{P}_\mathrm{i} = \hbar\vec{k}_\mathrm{f} + \vec{P}_\mathrm{f} \ , \qquad \hbar\omega_1(\vec{k}_\mathrm{i}) + \frac{\vec{P}_\mathrm{i}^2}{2M} = \hbar\omega_2(\vec{k}_\mathrm{f}) + \frac{\vec{P}_\mathrm{f}^2}{2M} \ , \tag{31.146}$$

où $\vec{P}_\mathrm{i}$ est l'impulsion des projectiles, cependant que $\vec{k}_{\mathrm{i,\,f}}$ et $\omega_j(\vec{k}_{\mathrm{i,\,f}})$ sont relatifs à une certaine excitation élémentaire de la cible, deux lois de dispersion $\omega_{1,\,2}$ pouvant être concernées (transition *bande à bande*) ou une seule (*transition intra-bande*). Ces deux égalités permettent de reconstituer point à point les relations de dispersion : la mesure de $\Delta E \stackrel{\text{déf}}{=} \frac{1}{2M}(\vec{P}_\mathrm{i}^2 - \vec{P}_\mathrm{f}^2)$ corrélée à celle de $\Delta P \stackrel{\text{déf}}{=} \vec{P}_\mathrm{i} - \vec{P}_\mathrm{f}$, permet de positionner le point $\omega_2(\vec{k}_\mathrm{f})$ par rapport au point $\omega_1(\vec{k}_\mathrm{i})$. Dans le cas usuel d'une excitation à partir du vide vers une branche $\omega(\vec{k})$, les deux équations de conservation conduisent à l'égalité unique :

$$\boxed{\omega\big(\hbar^{-1}(\vec{P}_\mathrm{i} - \vec{P}_\mathrm{f})\big) = \frac{1}{2M\hbar}(\vec{P}_\mathrm{i}^2 - \vec{P}_\mathrm{f}^2)} \tag{31.147}$$

les mesures du transfert de moment et de la perte d'énergie des neutrons fournissent[46] la fonction $\omega(\vec{k})$ cherchée.

La méthode dont le principe vient d'être décrit est une vision très simplifiée de toute situation réelle, et sa mise en œuvre est compliquée notamment par l'existence des polarisations ε_λ, dont les directions ne sont simples vis-à-vis de $\vec{k}$ que pour des cristaux de haute symétrie ponctuelle, ou pour certaines directions privilégiées de $\vec{k}$.

En outre, tout ce qui a été dit l'a été dans le cadre exclusif de l'approximation harmonique, de sorte que la comparaison avec une expérience n'est pas toujours aisée. En présence d'anharmonicité, les phonons ne sont plus des excitations propres (stationnaires), ce qui se traduit par l'existence d'interactions effectives entre eux (les phonons ne forment d'ailleurs plus un gaz parfait), que l'on peut se figurer comme donnant lieu à des collisions, évoquant celles entre atomes d'un gaz. De surcroît, comme le nombre de phonons n'est pas fixé[47], l'existence d'anharmonicités donne à ceux-ci une *durée de vie* finie. En définitive, au lieu d'avoir une loi de dispersion $\omega(\vec{k}\,)$ au sens strict, on a une relation *complexe*, du genre :

$$\tilde{\omega}(\vec{k}\,) = \omega(\vec{k}\,) + \Delta(\vec{k}\,) - \mathrm{i}\Gamma(\vec{k}\,) \ , \tag{31.148}$$

qui est typiquement le produit d'une théorie introduisant des *quasi-particules* au même titre que les quasi-particules brièvement présentées au début du chapitre 30 (voir p. 1364). Noter aussi que l'écriture (31.148) est en tout point analogue (pour un $\vec{k}$ donné) à l'énergie

[46] De toute évidence, la même procédure – éventuellement avec un faisceau de neutrons polarisés – fonctionne pour toute excitation élémentaire, (magnons, excitons,...).

[47] Il s'agit typiquement, comme pour le champ électromagnétique, d'un système *grand-canonique*.

effective, dans l'approximation du pôle, d'un atome couplé au champ électromagnétique (chapitre 25, sous-section 25.2, p.1167). Δ est un déplacement de pulsation (correction d'énergie, analogue du *Lamb shift*), cependant que Γ^{-1} donne la durée de vie du phonon. La partie imaginaire transforme les relations strictes de conservation (donnant des pics de Dirac) en *résonances* possédant une largeur finie. Toutes ces complications ne sont réellement sans importance qu'à très basse température : dès que la température monte, les vibrations prennent de l'ampleur et on ne peut plus omettre les effets anharmoniques.

31.9 Exercices et problèmes

31.9.1 Vibration d'une chaîne monoatomique 1D

Soit un réseau de Bravais unidimensionnel de paramètre a, dont le motif est un atome de masse M. Chaque atome effectue des petites vibrations le long du réseau dont l'amplitude est notée u_n ; à l'instant t, la position de l'atome situé près du site d'indice n est $x_n(t) = na + u_n(t)$. Le cas échéant, on utilisera des conditions cycliques en identifiant $N+1$ et 1.

Dans l'hypothèse des petites vibrations et d'interactions à courte portée, l'énergie potentielle de vibration est :

$$U_{\text{h}} = \frac{1}{2} K \sum_{n \in \mathbb{Z}} (u_{n+1} - u_n)^2 \ . \tag{31.149}$$

Quel est le signe de K et pourquoi ? Dans toute la suite, on pose $\omega_0 = \sqrt{K/M}$.

Équations du mouvement

1. Écrire les équations du mouvement (classiques) pour les déplacements u_n et les interpréter.
2. Pour résoudre ces équations, on pose $u_n(t) = \int_{-\infty}^{+\infty} U_n(\omega)\, \mathrm{e}^{-\mathrm{i}\omega t}\, \mathrm{d}\omega$. Écrire les équations satisfaites par les amplitudes $U_n(\omega)$.
3. Les solutions du système obtenu plus haut sont de la forme $U_n(\omega) = C\, \mathrm{e}^{\mathrm{i}nka}$, où ω est fonction de k. Déduire des résultats précédents la relation entre ω et k.
4. Suivant l'un des résultats précédents, il faut $\omega = +\omega(k)$ ou $\omega = -\omega(k)$, où $\omega(k)$ est une certaine fonction positive à préciser. Il en résulte que les déplacements peuvent être décomposés comme suit :

$$u_n(t) = \sum_k \mathrm{e}^{\mathrm{i}nka} \left(\xi_{k+} \mathrm{e}^{\mathrm{i}\omega(k)t} + \xi_{k-} \mathrm{e}^{-\mathrm{i}\omega(k)t}\right) \ , \tag{31.150}$$

où la somme sur k court sur les valeurs discrétisées situées, par exemple, dans la première zone de Brillouin. Donner les développements des déplacements initiaux $u_n(0)$ et des vitesses initiales correspondantes $\dot{u}_n(0)$.

5. Dans toute la suite, on suppose qu'à l'instant initial chaque atome est écarté de sa position d'équilibre sans vitesse initiale. Dans cette hypothèse, que devient l'expression de $u_n(t)$?

6. En utilisant les relations de Fourier[48], exprimer $u_n(t)$ en combinaison linéaire des déplacements initiaux $u_n(0)$.

7. Dans la limite $N \gg 1$, toutes les sommes discrètes sur k peuvent être remplacées par des intégrales. Écrire la forme finale de $u_n(t)$ qui résulte de cette substitution.

8. Établir l'égalité (31.52).

Étude à grande échelle

Dans cette partie, on suppose que le champ des déplacements varie très lentement à l'échelle a ($k \ll 1/a$) et on pose $a\omega_0 = c$.

1. Dans cette situation, que devient la loi de dispersion $\omega(k)$?

2. Quelle est la vitesse de groupe v_g ? Comment se compare-t-elle à la vitesse de phase v_ϕ ?

3. Pour les champs lentement variables, il est licite de transformer les expressions obtenues précédemment en prenant la limite continue d'espace – ce qui revient à faire tendre le paramètre a vers zéro. On posera :

$$\lim_{a \to 0,\, n \to +\infty} na = x \ , \qquad \lim_{a \to 0,\, n \to +\infty} u_n(t) = u(x,\, t) \ . \tag{31.152}$$

   Montrer que $u(x,\, t)$ se met alors sous la forme :

$$u(x,\, t) = \int_{-\infty}^{+\infty} \mathrm{d}x'\, G(x - x',\, t)\, u(x',\, 0) \tag{31.153}$$

   et donner l'expression intégrale de la fonction G.

4. En déduire l'équation de propagation satisfaite par le champ des déplacements $u(x,\, t)$. Connaissant $u(x,\, 0)$, donner l'expression formelle de $u(x,\, t)$ et commenter. Illustrer ces résultats par un schéma représentant la variation spatiale de u à $t = 0$ et à un instant quelconque t.

---

[48] Pour des variables conjuguées de Fourier, on a :

$$\alpha_n = \sum_k A_k\, \mathrm{e}^{\mathrm{i}kna} \quad \Longleftrightarrow \quad A_k = \frac{1}{N} \sum_n \alpha_n\, \mathrm{e}^{-\mathrm{i}kna} \ . \tag{31.151}$$

Retour au réseau discret

On suppose maintenant que le champ des déplacements varie vite à l'échelle a ; dans ces conditions, il convient de garder la loi de dispersion trouvée pour le réseau et l'expression première des déplacements $u_n(t)$.

1. Tracer le graphe de la variation de la vitesse de groupe, v_{g}, en fonction de k.
2. Donner l'expression approchée de v_{g} lorsque k est en bord de zone.
3. On suppose que les écarts initiaux sont répartis suivant $u_n(0) = (-1)^n\, u_0$, où u_0 est donné, ce qui définit une structure alternée présentant un ordre à longue distance. Trouver $u_n(t)$ sachant que :

$$\lim_{N \to +\infty} \sum_n \mathrm{e}^{inqa} = 2\pi\,\delta(qa) \qquad (-\pi/a \leq q \leq +\pi/a) \tag{31.154}$$

Que devient l'ordre initial au cours du temps ?

31.9.2 Vibration d'un réseau unidimensionnel de dimères

Soit un réseau unidimensionnel de pas[49] a, dont le motif est une molécule diatomique homonucléaire formée de deux atomes de masse M, séparés (à l'équilibre) de la distance $d < a$ (réseau de dimères, voir fig. 31.5) ; les constantes de raideur sont maintenant différentes : K_{i} pour le ressort entre les deux atomes séparés de d (vibration intramoléculaire), K pour deux atomes appartenant à deux molécules adjacentes ; afin de fixer les idées, on suppose $K_{\text{i}} > K$, qui est la situation physique usuelle. On désigne par $x_{1,\,j}$ l'écart de l'atome qui vibre autour de ja, $x_{2,\,j}$ l'écart de celui qui est près de $ja + d$. On pose :

$$\omega_0^2 = \frac{K_{\text{i}} + K}{M}\ , \qquad \Omega_0^2 = \frac{K}{M}\ , \qquad \Omega_{\text{i}}^2 = \frac{K_{\text{i}}}{M}\ , \qquad \Omega_0^2 + \Omega_{\text{i}}^2 = 2\omega_0^2\ , \tag{31.155}$$

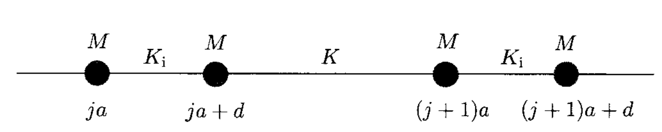


Figure 31.5: Positions d'équilibre du réseau de dimères.

1. Écrire les équations du mouvement.
2. En déduire le système déterminant les composantes de Fourier $X_{1,\,j}$ et $X_{2,\,j}$.

[49]Pour la comparaison avec les expressions obtenues avec celles trouvées pour le réseau à deux masses (section 31.3), se souvenir qu'alors la maille avait pour côté $2a$.

3. Trouver les relations de dispersion ; qualifier les deux branches obtenues.
4. Quel est le gap en bord de zone ?
5. Analyser les relations de dispersion pour $k \sim 0$ et $|k| \lesssim \frac{\pi}{2a}$.

31.9.3 Vibration d'un réseau avec des couplages harmoniques à longue portée

Dans toute la suite, on considère un réseau unidimensionnel de pas a formé de N atomes identiques de masse M. r_n désigne[50] la position instantanée du n^{e} atome :

$$r_n = R_n + u_n \ , \qquad R_n = na \ . \tag{31.156}$$

u_n est l'écart à la position d'équilibre. Afin de dépasser l'approximation ordinaire où l'on n'introduit que des couplages entre premiers voisins, l'énergie de vibration harmonique est écrite sous la forme suivante :

$$U_{\text{h}} = \sum_{n'=1}^{N} \sum_{m>0} \frac{1}{2} K_m (u_{n'} - u_{n'+m})^2 \ . \tag{31.157}$$

Il s'agit principalement de trouver la relation de dispersion de ce modèle et d'étudier ses particularités en fonction de la décroissance spatiale, lente ou rapide, des constantes de raideur K_m, supposées toutes positives. On pose $\omega_m^2 = \frac{K_m}{M}$.

1. Écrire le Lagrangien du réseau, $\mathcal{L}$.
2. Écrire l'équation de Lagrange pour u_n à l'aide de $\frac{\partial U_{\text{h}}}{\partial u_n}$.
3. Calculer explicitement $\frac{\partial U_{\text{h}}}{\partial u_n}$.
4. En déduire que l'équation du mouvement de u_n a la forme suivante :

$$M \ddot{u}_n = - \sum_{m>0} F_m \left(u_n, \, u_{n+m}, \, u_{n-m}\right) \ , \tag{31.158}$$

où F_m est une fonction linéaire de ses arguments dont on précisera l'expression.

5. Afin d'obtenir la solution de l'équation (31.158), on pose $u_n = C \, \mathrm{e}^{-\mathrm{i}\omega t} \, \mathrm{e}^{\mathrm{i}kna}$, où C est une constante. Combien vaut la partie imaginaire de k ?
6. En déduire que la pulsation ω est donnée par :

$$\omega = \pm \, \omega(k) \ , \qquad \omega(k) = 2 \left[\sum_{m>0} \lambda_m(k) \right]^{1/2} \ . \tag{31.159}$$

Donner l'expression des $\lambda_m(k)$.

[50] Comme on l'a vu, r_n (ou u_n) est indifféremment une variable classique ou la valeur moyenne quantique de l'opérateur position dans la description de Heisenberg.

7. Dans la limite des déplacements lentement variables à l'échelle du réseau (approximation des grandes longueurs d'onde), montrer que la relation de dispersion est à peu près donnée par :
$$\omega(k) \simeq |k|c \ , \tag{31.160}$$
à condition toutefois que la série $S \equiv \sum_m m^2\omega_m^2$ soit convergente. Donner l'expression de la vitesse du son c.

8. Si l'on admet que K_m a une décroissance algébrique du genre :
$$K_m = \frac{K}{m^p} \ , \tag{31.161}$$
quelle est la condition sur l'exposant p pour que le résultat (31.160) ait un sens ?

9. On suppose maintenant $1 < p < 3$ (couplages harmoniques à très longue portée). Dès lors, il n'est plus sensé de remplacer le sinus par son argument ; toutefois, la limite des grandes longueurs d'onde autorise à substituer la somme $\sum_m$ par une intégrale sur l'abscisse continue x :
$$\sum_{m>0} K_m \sin^2\frac{kma}{2} \quad \rightarrow \quad Ka^q \int_0^{+\infty} \mathrm{d}x\, x^{-p} \sin^2\frac{kx}{2} \ . \tag{31.162}$$
Combien vaut l'exposant q ?

10. Effectuer le changement de variable convenable pour obtenir l'expression suivante :
$$\omega(k) = 2\, A_p\sqrt{K/M}\, (ka)^{(p-1)/2} \ , \tag{31.163}$$
où A_p est un nombre. À quelle condition sur p ce nombre est-il fini ? Donner l'allure de la loi de dispersion dans le voisinage de $k = 0$. Peut-on dans ce cas définir une vitesse du son ?

31.9.4 Instabilité de Peierls

Il s'agit de montrer qu'un gaz d'électrons sur un réseau unidimensionnel déformable est instable et a tendance à induire une surstructure cristalline. Dans le cas considéré ici (un électron par atome), on assiste à une dimérisation du réseau, l'ion initialement situé en $x_n^{(0)} = na$ quand le réseau est rigide étant finalement décalé d'une petite quantité $u_n = (-1)^n u$, où u est un déplacement constant ($x_n = na + (-1)^n u$). Un exemple physique de cette situation est constitué par certaines longues chaînes carbonées, comme le polyacétylène, rencontré dans le chapitre 28, sous-section 28.2.5 (voir aussi problème 28.4.5).

La partie 1 est une analyse qualitative du phénomène ; la partie 2 en fournit une description détaillée sur la base d'un modèle simplifié.

1. (a) Dans toute la suite, les électrons sont supposés indépendants, de sorte que l'on raisonne dans le cadre d'une théorie à une particule ; par ailleurs, le cas échéant, on invoquera les conditions cycliques de Born - von Kármán relatives à une chaîne bouclée sur elle-même et contenant N sites.

   Soit $\varepsilon(k)$ la loi de dispersion de la plus haute bande occupée, présentant un minimum en centre de zone et recevant les N électrons de valence atomiques. Quel est son taux de remplissage ? Combien vaut k_F, nombre d'onde de Fermi associé aux plus hauts niveaux occupés ?

   (b) Le réseau d'ions n'étant pas rigide, expliquer pourquoi l'énergie du gaz d'électrons, E_e, est abaissée si les ions sont déplacés de u_n.

   (c) L'abaissement de E_e consécutif à la déformation de réseau est en compétition avec l'augmentation d'énergie élastique de réseau, E_harm. Dans toute la suite, on suppose le réseau pas trop mou pour qu'un compromis énergétique puisse être effectivement réalisé ; quelle est la conséquence de la dimérisation sur les propriétés de conduction électrique du solide ?

2. Le Hamiltonien d'un électron sur le réseau est écrit comme suit :

$$H = -\sum_n t_{nn+1}(|\phi_n\rangle\langle\phi_{n+1}| + |\phi_{n+1}\rangle\langle\phi_n|) \ . \tag{31.164}$$

$|\phi_n\rangle$ est une fonction bien localisée centrée sur le site d'abscisse x_n ; les $|\phi_n\rangle$ sont orthonormalisées :

$$\langle\phi_n|\phi_n\rangle = \delta_{nn'} \ . \tag{31.165}$$

t_{nn+1} est l'amplitude (réelle positive) permettant le saut de l'électron d'un site vers les premiers voisins et dépend de la distance $x_n - x_{n+1}$ entre deux sites ; en partant de l'idée que le déplacement u est toujours très petit devant a, on se borne à une variation linéaire et on pose :

$$t_{nn+1} = t_0 - \alpha(x_n - x_{n+1}) \ . \tag{31.166}$$

t_0 et α sont des quantités positives, dont les ordres de grandeur sont : $t_0 \sim 1$ eV, $\alpha \sim 1$ eV/Å.

   (a) On décompose le Hamiltonien (31.164) comme suit :

$$H = -t_0\sum_n(|\phi_n\rangle\langle\phi_{n+1}| + |\phi_{n+1}\rangle\langle\phi_n|) + V \ . \tag{31.167}$$

   Donner l'expression de V.

   (b) Soit l'état de Bloch normalisé $|\psi_k\rangle$:

$$|\psi_k\rangle = \frac{1}{\sqrt{N}}\sum_n \mathrm{e}^{ikna}\,|\phi_n\rangle \qquad (k = \text{entier}\ \frac{2\pi}{N}) \ . \tag{31.168}$$

   Montrer que $|\psi_k\rangle$ est un état propre de H_0 ; quelle est l'énergie propre correspondante, $\varepsilon^{(0)}(k)$? La tracer sur l'intervalle $[-2k_\mathrm{F}, +2k_\mathrm{F}]$.

(c) Dessiner le schéma en zones réduites sur l'intervalle $[-k_{\rm F},\, +k_{\rm F}]$. Dans toute la suite, on pose, $\forall k \in [-k_{\rm F},\, +k_{\rm F}]$:

$$\varepsilon_-^{(0)}(k) = \varepsilon^{(0)}(k) \ , \qquad \varepsilon_+^{(0)}(k) = \varepsilon^{(0)}(k + 2k_{\rm F}) \ , \tag{31.169}$$

$$|\psi_{k,-}\rangle = |\psi_k\rangle \ , \qquad |\psi_{k,+}\rangle = |\psi_{k+2k_{\rm F}}\rangle \ . \tag{31.170}$$

(d) Ecrire le développement de $|\psi_{k,+}\rangle$ sur les $|\phi_n\rangle$.

(e) Combien vaut $V|\psi_{k,-}\rangle$? En déduire que H est une somme d'opérateurs H_k définis chacun dans le sous-espace engendré par $|\psi_{k,-}\rangle$ et $|\psi_{k,+}\rangle$. Ecrire la matrice représentant H_k sur cette base.

(f) En déduire que les valeurs propres de H sont $\pm\varepsilon(k)$ où :

$$\varepsilon(k) = -2t_0\sqrt{1 - \Lambda\sin^2 ka} \ ; \tag{31.171}$$

préciser le paramètre Λ. Représenter la variation de $\varepsilon(k)$ sur $[-k_{\rm F},\, +k_{\rm F}]$; illustrer par un schéma la répartition des électrons dans l'état fondamental (mer de Fermi).

(g) Donner l'expression de $E_{\rm e}$, énergie du gaz d'électrons dans l'état fondamental. Dans la limite $N \to +\infty$, exprimer $\varepsilon_{\rm e} \equiv E_{\rm e}/N$ à l'aide de la fonction $\mathbf{E}(\Lambda)$, fonction elliptique complète de deuxième espèce, définie comme :

$$\mathbf{E}(\Lambda) \overset{\text{déf}}{=} \int_0^{\pi/2} \sqrt{1 - \Lambda\sin^2 x}\,\mathrm{d}x \qquad (\Lambda \le 1) \ . \tag{31.172}$$

(h) L'énergie élastique de déformation est écrite sous la forme la plus simple possible :

$$E_{\rm harm} = \sum_n \frac{1}{2}K(x_n - x_{n+1})^2 \ . \tag{31.173}$$

Combien vaut $\varepsilon_{\rm harm} \equiv E_{\rm harm}/N$?

(i) Au total, l'énergie par particule, ε, est telle que :

$$\tilde{\varepsilon} \equiv \frac{\varepsilon}{2t_0} = -\frac{2}{\pi}\mathbf{E}(\Lambda) + \frac{K}{t_0}u^2 \ . \tag{31.174}$$

Comme $u/a \ll 1$, il suffit de connaître le développement de $\mathbf{E}(\Lambda)$ dans le voisinage à gauche de $\Lambda = 1$:

$$\mathbf{E}(\Lambda) \simeq 1 + \frac{1}{4}(1-\Lambda)\left(\ln\frac{16}{1-\Lambda} - 1\right) \ . \tag{31.175}$$

En posant :

$$\xi = \frac{K}{\alpha^2 t_0} \qquad y = \frac{\alpha u}{t_0} \ , \tag{31.176}$$

$\tilde{\varepsilon}$ est une certaine fonction $f(y;\,\xi)$. Donner l'allure de f en fonction de y à ξ donné.

(j) f est minimum en un certain y_0. Montrer que $y_0 \ll 1$ est de la forme :

$$y_0 \simeq C\,\mathrm{e}^{-c\xi} \tag{31.177}$$

où C et c sont des constantes.

(k) Calculer les valeurs numériques de u_0 et de l'abaissement d'énergie totale par électron résultant de la dimérisation du solide, avec $t_0 = 2\,\mathrm{eV}$, $\alpha = 2\,\mathrm{eV/Å}$, $K = 20\,\mathrm{eV/Å}^2$.

31.9.5 Identité de Bloch

1. Démontrer l'identité de Bloch, (31.130).

2. En prenant $\gamma = \mu = \mathrm{i}\alpha$, et en utilisant le noyau de l'opérateur densité canonique de l'oscillateur harmonique (Tome I, éq. (16.157)), retrouver le résultat précédent.

3. Qu'en est-il si $\gamma = -\mu = \mathrm{i}\alpha$?

physique

Chapitre 32

Notions de transport dans les solides

Ce dernier chapitre est consacré au problème du transport dans un solide ordonné. On présente le modèle semi-classique permettant de décrire convenablement les situations les plus courantes et d'obtenir l'expression de la conductivité électrique σ, mettant ainsi en évidence l'importance de la masse effective. L'ouvrage se termine par une introduction à la supraconductivité, donnant l'occasion d'un exposé en raccourci de l'une des plus belles théories construites dans le cadre quantique : la théorie BCS.

32.1 Généralités

Dans un milieu, quel qu'il soit, l'existence d'un transport (de matière, de charge, de chaleur, etc.) résulte d'un déséquilibre conduisant à l'apparition de courants. Il peut s'agir de courants transitoires, quand on part d'une situation hors d'équilibre et que le système, abandonné à lui-même, relaxe vers l'équilibre thermodynamique. Par exemple, une impulsion thermique ponctuelle en temps et en espace provoque un gradient de température qui induit un courant de chaleur transitoire, disparaissant lorsque l'équilibre thermodynamique est retrouvé. Il peut aussi s'agir de courants permanents, provoqués par un déséquilibre maintenu de l'extérieur : quand on soumet un métal – volume équipotentiel à l'équilibre – à une différence de potentiel imposée de l'extérieur[1], il y apparaît un courant électrique permanent, donné par la loi d'Ohm.

D'une façon générale, la description fine d'une situation hors d'équilibre est d'une grande difficulté et passe par la manipulation d'équations contenant à la fois la dynamique

[1]se traduisant en pratique par des conditions aux limites.

(mouvement en terme de forces) et la relaxation (retour à l'équilibre), seulement rendue possible s'il y a une dissipation d'énergie.

La nécessité de la dissipation apparaît immédiatement grâce à un modèle classique et très simple qui, de plus, servira de comparaison au résultat du calcul quantique de la conductivité σ d'un métal[2]. Il s'agit du modèle de Drude (1905), qui représente un métal essentiellement comme un gaz d'électrons libres soumis à une force de freinage fluide, et montre à peu de frais un fait majeur : en l'absence de dissipation, il n'y aurait pas de loi d'Ohm, c'est-à-dire pas de proportionnalité[3] entre le courant électrique $\vec{j}$ et le champ appliqué $\vec{\mathcal{E}}$:

$$\boxed{\vec{j} = \sigma\vec{\mathcal{E}}} \tag{32.1}$$

le courant ayant l'expression ordinaire, $\vec{j} = n_\mathrm{e}e\vec{v}$. S'il n'y avait pas de freinage, l'application du champ donnerait aux électrons un mouvement uniformément accéléré et donc une vitesse augmentant linéairement en temps : le courant ne saurait être constant ! Il y a donc forcément un phénomène venant contrecarrer l'augmentation de la vitesse et, en l'absence d'autre forces contrôlées, la seule possibilité est un *freinage* d'origine microscopique ouvrant la possibilité d'états permanents où d'une part la vitesse est effectivement constante, où d'autre part les *conditions initiales* sont oubliées. Restant dans une vision purement classique, on imagine un électron accéléré par le champ mais entrant en collision avec les ions du métal et effectuant, finalement, une sorte de *marche au hasard* dans le réseau.

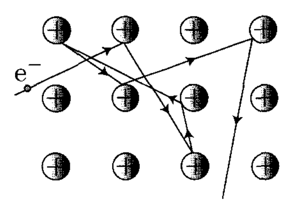


Figure 32.1: Marche au hasard d'un électron dans un métal (Drude).

La façon la plus simple de traduire le freinage est l'adjonction *à la main* d'une force de freinage fluide juste proportionnelle à la vitesse, d'où l'équation dynamique :

$$m\,\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = e\vec{\mathcal{E}} - \alpha\,\vec{v} \ , \tag{32.2}$$

qui possède une solution à vitesse constante donnée par :

$$\vec{v} = \frac{e\tau}{m}\,\vec{\mathcal{E}} \ ; \tag{32.3}$$

[2]D'ailleurs, l'expression de la conductivité selon Drude, formule (32.4), est "presque" exacte : on verra par la suite que, souvent, la modification essentielle (et importante) consiste à remplacer la masse m (masse nue) de l'électron par sa masse effective – voir (32.64).

[3]Le plus souvent, la relation de proportionnalité entre courant et champ n'est pas simplement scalaire puisque la conductivité peut être anisotrope (exemple : le graphite). À ce stade, on ne rentre pas dans ce genre de finesse.

ce régime prévaut pour $t \gg \tau$, temps d'extinction du transitoire incoporant la vitesse initiale, montrant que τ est un temps de relaxation. Physiquement, cette relaxation provient de collisions inélastiques[4] au sens large (collisions avec les ions du réseau, avec des défauts, etc.) ; le champ accélère l'électron, mais chaque collision réduit cette vitesse. Au total, la vitesse sature en moyenne à une valeur finie constante, indépendante de la condition initiale[5], et proportionnelle au champ appliqué. Revenant à l'expression du courant et y injectant cette vitesse limite, on obtient l'expression de la conductivité statique (fréquence nulle), σ_0 :

$$\boxed{\sigma_0 = \frac{n_\mathrm{e} e^2 \tau}{m}} \tag{32.4}$$

Ce résultat se généralise sans peine au cas d'un champ variable $\vec{\mathcal{E}} = \vec{\mathcal{E}}_0\,\mathrm{e}^{\mathrm{i}\omega t}$. Dans ces conditions, la composante de Fourier $V(\omega)$ de la vitesse vaut $\frac{\tau}{m(1+\mathrm{i}\omega\tau)} e\vec{\mathcal{E}}_0$; celle du courant, $\vec{J}(\omega)$, est alors $n_\mathrm{e} e\vec{V}(\omega)$, d'où la conductivité en fréquence $\sigma(\omega)$:

$$\boxed{\sigma(\omega) = \frac{n_\mathrm{e} e^2 \tau}{m(1+\mathrm{i}\omega\tau)} \equiv \frac{\sigma_0}{1+\mathrm{i}\omega\tau}} \tag{32.5}$$

La conductivité tend vers zéro aux très hautes fréquences, un résultat qui se comprend bien physiquement : si la force électrique change de sens trop rapidement, l'électron n'a pas le temps de suivre le mouvement et le courant devient nul.

Le modèle de Drude ne vaut que pour un métal, son ingrédient essentiel étant une densité n_e de porteurs libres (ou quasi-libres), qui n'existe ni dans un semi-conducteur, ni dans un isolant. Par ailleurs, la relation (32.4) peut être inversée :

$$\tau = \frac{m}{n_\mathrm{e} e^2}\,\sigma_0 \equiv \frac{m}{n_\mathrm{e} e^2}\,\frac{1}{\rho} \ , \tag{32.6}$$

où $\rho = 1/\sigma_0$ est la résistivité usuelle, afin d'exprimer le temps de relaxation en termes de quantités facilement mesurables. On peut ainsi trouver que τ est, à la température ambiante, de l'ordre de 10^{-14} à 10^{-15} s. Ce résultat permet de vérifier *a posteriori* la cohérence relative du modèle de Drude ; en effet, τ doit être du même ordre de grandeur que l'intervalle de temps entre deux collisions puisque quelques collisions suffisent pour effacer la mémoire de toute condition initiale, les diffuseurs étant indépendants les uns des autres. Connaissant l'ordre de grandeur de τ, le choix d'une bonne vitesse permet de fabriquer une longueur jouant le rôle d'un libre parcours moyen, qui doit être de l'ordre de grandeur d'un paramètre de maille pour que le modèle soit cohérent ; partant de $m\vec{v}_\mathrm{th}^{\,2}/2 = (3/2)k_\mathrm{B}T$, on trouve, à l'ambiante $v_\mathrm{th} \sim 10^5$ m/s, ce qui donne $v_\mathrm{th}\tau \sim 5$ Å. Cette distance est comparable à la distance entre ions du métal et, interprétée comme

[4] Il y a bien de l'inélasticité quelque part, puisque (32.2) entraîne que l'énergie mécanique de l'électron n'est pas conservée.

[5] Cette amnésie est garantie pour un système linéaire, dont l'état d'équilibre, toujours accessible grâce à la relaxation, est défini en soi par la donnée de conditions externes excluant la nécessité d'information inaccessible, la vitesse initiale de l'électron par exemple. L'expression au second membre de (32.3) est aussi très exactement la vitesse moyenne (au sens statistique) quand la vitesse initiale $\vec{v}_0$ est distribuée uniformément au hasard, hypothèse fort raisonnable physiquement.

un libre parcours moyen, justifie grossièrement l'image de la dissipation d'énergie par collisions successives des porteurs de charge sur les ions, étant entendu que, en tout état de cause, le modèle de Drude, ne saurait être trop pris au sérieux, exposé ici dans le seul but de mettre en avant le rôle crucial de la dissipation pour expliquer une loi aussi élémentaire que la loi d'Ohm.

Les deux aspects – dynamique et dissipation – devant être inclus, le problème est d'une redoutable complexité. Classiquement, il faudrait écrire une équation de Boltzmann pour la densité dans l'espace des phases, $\rho(\{\vec{r}\}, \{\vec{p}\}, t)$, qui est une équation aux dérivées partielles, intégrale et non-linéaire – tout pour plaire. S'agissant ici d'aborder le calcul de la conductivité d'un métal, il faut bien évidemment délaisser le cadre classique : la discussion de la distinction métal/isolant a clairement mis en lumière le fait que seule la Mécanique Quantique peut en fournir une interprétation cohérente, de sorte qu'à la complexité inhérente s'ajoute la nécessité d'en donner une description quantique.

On s'en tiendra ici à une description élémentaire de la conductivité électrique, où les indispensables phénomènes dissipatifs[6] sont introduits *à la main*, sur la base d'arguments plausibles. Quant au mouvement purement *mécanique*, sa description quantique est en soi fort complexe ; on se bornera à utiliser le modèle dit *semi-classique* dont la justification précise, longue et difficile, permet d'en comprendre la validité.

L'exposé suivant procède donc en deux étapes principales :

1. dans un premier temps (section 32.2), on construit une description purement mécanique en termes de forces, incorporant l'essentiel des caractéristiques du mouvement quantique d'une particule dans un potentiel périodique.
2. Dans un deuxième temps (section 32.4), la dissipation est introduite de façon phénoménologique, sur la base d'arguments acceptables par simple bon sens.

32.2 Le modèle semi-classique

Commençons par réinterpréter le résultat donnant la vitesse moyenne d'un électron dans l'état de Bloch $\psi_{\vec{k}_0 n}$, associé à la valeur propre $\varepsilon_n(\vec{k}_0)$ du Hamiltonien à une particule (chapitre 30, sous-section 30.3.3, éq. (30.64)) :

$$\vec{v}_n(\vec{k}_0) \equiv \langle \psi_{n\vec{k}_0} | \frac{\vec{p}}{m} | \psi_{n\vec{k}_0} \rangle = \hbar^{-1} \left(\frac{\partial \varepsilon_n}{\partial \vec{k}} \right)_{\vec{k}_0} , \tag{32.7}$$

[6]Il peut s'agir de collisions entre les porteurs de charge (les électrons usuellement) et des défauts du cristal (lacunes, impuretés, etc.), toujours présents en pratique. Conceptuellement, on peut bien sûr imaginer un cristal parfait, mais même dans ce cas, la possibilité de dégradation existe encore, par collision entre les électrons et les phonons.

qui n'est rien d'autre que la vitesse de groupe associée à la loi de dispersion $\varepsilon_n(\vec{k})$. Soit un paquet d'ondes formé sur des états de Bloch relatifs à *une* bande d'indice n donné :

$$\Psi(\vec{r},\, t) = \sum_{\vec{k}} A(\vec{k})\, \psi_{\vec{k}}(\vec{r})\, \mathrm{e}^{-\mathrm{i}\omega(\vec{k})t} \quad , \qquad \omega(\vec{k}) = \hbar^{-1}\varepsilon(\vec{k}) \ . \tag{32.8}$$

L'amplitude $A(\vec{k})$ est caractérisée par une certaine largeur en $\vec{k}$, Δk ; désignant par $\vec{k}_0$ le vecteur où A est maximum, on a par construction $A(\vec{k}) \simeq 0$ si $\|\vec{k} - \vec{k}_0\| > \Delta k$. Si Δk est très petit par rapport à la taille $\frac{\pi}{a}$ de la première zone de Brillouin, la vitesse de groupe de ce paquet d'ondes n'est autre que la vitesse donnée par (32.7) – résultat d'ailleurs quasi-évident : avec un tout petit Δk, l'état $\Psi(\vec{r},\, t)$ est très voisin de l'état de Bloch (stationnaire) caractérisé par $\vec{k}_0$. Par les inégalités de Heisenberg, ce paquet d'ondes a une largeur spatiale Δr d'ordre $\frac{1}{\Delta k}$ qui est par hypothèse très grand devant a : il s'étend donc sur un très grand nombre de mailles primitives de $\mathcal{B}$. Dans la suite, on raisonnera essentiellement avec pour seul ingrédient la loi de dispersion d'une bande donnée, $\varepsilon_n(\vec{k})$, où l'argument $\vec{k}$ de la fonction doit être compris comme la valeur de la quasi-impulsion correspondant au maximum de l'amplitude du paquet d'ondes, par ailleurs supposé très étroit à l'échelle a^{-1}, dans l'espace réciproque.

Par ailleurs, en vertu du théorème d'Ehrenfest (Tome I, sous-section 14.1.3), un paquet d'ondes soumis à un champ extérieur variant lentement relativement aux dimensions du paquet d'ondes a un mouvement quasi-classique, son centre suivant en gros la trajectoire classique. Pour cette raison, et désignant par l l'échelle de variation spatiale typique du champ extérieur appliqué, on admettra au total dans toute la suite que les inégalités suivantes sont satisfaites :

$$\boxed{a \ll \frac{1}{\Delta k} \ll l} \tag{32.9}$$

La longueur a étant atomique et l macroscopique, les inégalités (32.9) sont faciles à satisfaire[7]. Ceci étant posé, on sera autorisé dans la suite à convenir de parler d'un électron *de vitesse et de position déterminées* pour dire que les incertitudes sur les grandeurs $\hbar\vec{k}$ et $\vec{r}$ sont l'une très petite par rapport à $\frac{\hbar}{a}$, l'autre très petite par rapport à la longueur macroscopique l. Dans la suite, on notera $\vec{r}$ la "position" de l'électron, étant entendu qu'il s'agit de sa position moyenne (au sens quantique), très dispersée à l'échelle atomique mais très précise à l'échelle l.

Le modèle de transport semi-classique repose entièrement sur le seul ingrédient $\varepsilon_n(\vec{k})$: il présuppose la seule connaissance préalable de ces énergies de bande (sans rien dire sur la façon de les obtenir, mais ce n'est pas son objet), et donne des prescriptions permettant de trouver les propriétés de transport (réponse du système à des champs, ou à des gradients de température) ; inversement, et sa validité étant avérée dans le cas considéré, la mesure des propriétés de transport permet d'inférer des caractéristiques des énergies de bande. Ses prescriptions se résument en l'énoncé de trois affirmations :

[7] On sait que les paquets d'ondes s'élargissent au cours du temps : le temps passant, la largeur Δr peut donc se rapprocher de l. On admettra que tout intervalle de temps considéré est suffisamment petit pour que les inégalités (32.9) soient toujours satisfaites – gardant à l'esprit que, de toute façon, le modèle ici développé ne vise à décrire que le mouvement purement mécanique, c'est-à-dire entre deux collisions.

1. l'index de bande n ne change pas (est une constante du mouvement). Ceci élimine la possibilité de transitions inter-bandes et suppose donc les champs externes petits par rapport aux champs microscopiques, en tout cas suffisamment faibles pour ne pas donner à un électron assez d'énergie pour franchir un *gap*. Avec les champs usuels, cette hypothèse est facilement satisfaite[8].

2. En présence d'un champ électromagnétique $(\vec{\mathcal{E}}(\vec{r}, t), \vec{\mathcal{B}}(\vec{r}, t))$, les équations (du mouvement) donnant le taux de variation de la "position" $\vec{r}$ et de la pseudo-impulsion $\hbar\vec{k}$ sont les suivantes[9] :

$$\boxed{\dot{\vec{r}} \stackrel{\text{déf}}{=} \vec{v}_n(\vec{k}) = \frac{1}{\hbar}\frac{\partial \varepsilon_n}{\partial \vec{k}}} \tag{32.10}$$

$$\boxed{\hbar\dot{\vec{k}} = e\left[\vec{\mathcal{E}}(\vec{r}, t) + \vec{v}_n(\vec{k}) \times \vec{\mathcal{B}}(\vec{r}, t)\right]} \tag{32.11}$$

La première égalité (définissant $\dot{\vec{r}}$) est visiblement nécessaire pour que ces équations aient un sens (il faut bien savoir ce qu'est la *vitesse* liée à la position $\vec{r}$).

3. Le remplissage des bandes s'effectue conformément au Principe de Pauli, en suivant la distribution de Fermi - Dirac :

$$f(\varepsilon_n(\vec{k})) = \frac{1}{\mathrm{e}^{\beta[\varepsilon_n(\vec{k})-\mu]}+1} \qquad (\beta = (k_\mathrm{B}T)^{-1}) \ . \tag{32.12}$$

Bien évidemment, comme les états de l'électron sont définis à une translation près dans $\tilde{\mathcal{B}}$, les deux états $\vec{k}$ et $\vec{k}+\vec{K}$ sont identiques : il ne peut exister deux électrons ayant le même indice de bande, n, le même spin et ne différant que par un vecteur $\vec{K}$. Cela étant, pour un nombre donné d'électrons (par unité de volume, par exemple), les états sont remplis, à température nulle, jusqu'à un certain niveau d'énergie ε_F.

Avant de poursuivre, quelques commentaires s'imposent. L'équation (32.10) traduit la propriété d'un paquet d'ondes ayant une faible dispersion en $\vec{k}$, et semble très naturelle ; on peut la considérer comme une équation de Hamilton, $\varepsilon_n(\vec{k})$ jouant le rôle d'un Hamiltonien. L'équation (32.11), en revanche, peut conduire à une méprise : en dépit des apparences, elle n'est pas la version quantique d'une équation de Hamilton, car le premier membre *n'est pas* l'impulsion de l'électron (ou, plus précisément, sa valeur moyenne dans un paquet d'ondes peu dispersé). Outre les éventuelles forces extérieures, l'électron est soumis au potentiel du réseau, complètement pris en compte (à noyaux fixés) par l'intervention de $\varepsilon_n(\vec{k})$. D'ailleurs, comme $\vec{k}$ est un bon nombre quantique, c'est une constante du mouvement en l'absence de forces extérieures, i.e. en l'absence de

[8] On se souvient des ordres de grandeurs trouvés lors de l'étude de l'effet Stark pour un atome : un champ de 10^4 V/cm est tout petit par rapport à un champ intra-atomique ($\sim 10^9$ V/cm).

[9] Avec l'hypothèse $\Delta k \ll a^{-1}$, on confond $\vec{k}$ et sa valeur moyenne, ce qui met finalement le paquet d'ondes (32.8) au second plan. L'équation (32.10) doit bien être considérée comme la *définition* de $\dot{\vec{r}}$, étant entendu que $\vec{r}$ désigne en fait la *moyenne* (quantique) de l'observable position dans un paquet d'ondes ayant les propriétés ci-dessus.

forces non incluses dans le Hamiltonien conduisant aux lois de dispersion exprimées à l'aide de $\vec{k}$. Physiquement, seule une force *extérieure* peut changer l'énergie $\varepsilon(\vec{k})$ d'un électron réputé être dans un état propre, ce qui ne peut se produire que si $\vec{k}$ varie d'une façon ou d'une autre. Le modèle semi-classique retient que la variation de $\vec{k}$ est donnée par (32.11).

La limite de validité des équations semi-classiques est difficile à cerner en détail. On se bornera à énoncer quelques faits d'évidence. Comme ce modèle interdit les transitions inter-bandes, toute énergie liée aux champs externes doit être très petite par rapport aux énergies de *gap*. Ces dernières sont d'autant plus grandes que le potentiel de réseau est grand ; autrement dit, une première condition de validité est que celui-ci doit être assez intense. Par ailleurs, avec les modules des champs, $\mathcal{E}$ et $\mathcal{B}$, on peut former deux énergies, $|e|\mathcal{E}a$ et $\hbar\,(|e|\mathcal{B}/m) \stackrel{\text{déf}}{=} \hbar\omega_{\text{c}}$; les deux conditions suivantes doivent donc être *a priori* vérifiées :

$$|e|\mathcal{E}a \ll E_{\text{gap}} \ , \qquad \hbar\omega_{\text{c}} \ll E_{\text{gap}} \ . \tag{32.13}$$

Ces inégalités laissent de côté une énergie pourtant importante, l'énergie de Fermi ε_{F}, qui doit apparaître dans les conditions de validité. Une analyse plus subtile montre que les bonnes conditions sont plus précisément :

$$\boxed{|e|\mathcal{E}a \ll \frac{E_{\text{gap}}^2}{\varepsilon_{\text{F}}}} \qquad \boxed{\hbar\omega_{\text{c}} \ll \frac{E_{\text{gap}}^2}{\varepsilon_{\text{F}}}} \tag{32.14}$$

La première condition (32.14) est usuellement satisfaite ; en prenant des valeurs typiques $E_{\text{gap}} \sim 0,1\,\text{eV}$ et[10] $\varepsilon_{\text{F}} \sim 1\,\text{eV}$, le rapport de droite vaut $\sim 10^{-2}\,\text{eV}$. Le premier membre vaut $\sim 10^{-10}\mathcal{E}\,\text{eV}$ et ne devient comparable au second que pour $\mathcal{E} \sim 10^8\,\text{V/m}$, une valeur gigantesque[11]. Au contraire, la deuxième condition (32.14) est violée pour des champs magnétiques que l'on peut réaliser facilement ($\sim 1\,\text{T}$) – on parle alors de claquage magnétique (*magnetic breakdown*).

Comme en général $\frac{E_{\text{gap}}}{\varepsilon_{\text{F}}}$ est petit devant 1, les conditions (32.14) sont plus nettement plus contraignantes que (32.13). Ceci ne doit pas surprendre : finalement, cette dernière condition ne s'appuie que sur la nécessité de la pertinence d'une seule bande, qui ne suffit pas à justifier les équations du mouvement (32.10) et (32.11).

Avec des champs externes variables dans le temps à la pulsation ω et dans l'espace à la longueur d'onde λ, il faut évidemment $\hbar\omega \ll E_{\text{gap}}$ et $\lambda \gg a$. Enfin, il convient d'avoir toujours à l'esprit que, à température finie, pour des raisons physiques évidentes, seuls les électrons situés au voisinage de l'énergie de Fermi dans un intervalle d'ordre $k_{\text{B}}T$ sont concernés par le transport et y participent effectivement.

[10] $1\,\text{eV} \simeq 12\,000\,\text{K}$: on sait que ε_{F} est typiquement de l'ordre de 10 000 à 20 000 K.

[11] Bien sûr, avec des champs électriques très intenses, on peut provoquer un phénomène de claquage (*electric breakdown*). Le champ est alors assez fort pour provoquer des passages d'une bande à l'autre ; l'énergie d'un électron peut alors augmenter vertigineusement, d'où le claquage conduisant à un plasma.

32.3 Conséquences du modèle semi-classique

Une première conséquence des hypothèses définissant le modèle semi-classique est l'inertie absolue de toute bande complète. Séparée en énergie de toutes les autres par un *gap* et pleine par hypothèse, il ne peut y avoir de mouvement électronique : dans cette bande, tous les états sont remplis, de sorte que tout "courant" entre états apparaissant dans un sens est forcément compensé par un courant identique mais dirigé en sens contraire – dont le mécanisme sera précisé ci-dessous[12]. Inversement, la conduction apparaît due entièrement aux bandes incomplètes, là où les électrons peuvent trouver des places libres, moyennant un peu d'énergie supplémentaire apportée par les champs externes.

32.3.1 Mouvement dans un champ électrique constant dans le temps

En présence d'un champ électrique constant, les équations semi-classiques deviennent :

$$\dot{\vec{r}} \equiv \vec{v}_n(\vec{k}) = \frac{1}{\hbar}\frac{\partial \varepsilon_n}{\partial \vec{k}} \ , \qquad \hbar\dot{\vec{k}} = e\,\vec{\mathcal{E}}(\vec{r}) \ , \tag{32.15}$$

d'où :

$$\boxed{\vec{k}(t) - \vec{k}(0) = \frac{e}{\hbar}\,\vec{\mathcal{E}}(\vec{r})\,t} \tag{32.16}$$

L'écart entre le vecteur d'onde à l'instant t et sa valeur de départ augmente donc linéairement avec le temps ; compte tenu des prescriptions antérieures permettant de ne considérer le modèle semi-classique que comme un prérequis du transport, cette variation dans le temps n'a de sens que si t reste petit devant le temps τ séparant deux collisions successives (ou en tout cas n'est pas plus grand que τ). Ceci permet de se faire une idée numérique de l'augmentation Δk de $\vec{k}$ sous l'effet du champ. Prenant $\mathcal{E} = 10\,\mathrm{V/cm}$ et $\tau = 10^{-14}\,\mathrm{s}$, on trouve $\Delta k \sim e\hbar^{-1}\mathcal{E}\tau \sim 10^2\,\mathrm{cm}^{-1}$, alors que la dimension linéaire de la zone est d'ordre $10^8\,\mathrm{cm}^{-1}$. Cette variation – la seule qui ait un sens avec la condition $t \lesssim \tau \sim 10^{-13}\,\mathrm{s}$ – est donc microscopique comparée à la taille de la zone. D'ailleurs, l'augmentation d'énergie correspondante, $\hbar^2(\Delta k)^2/(2m)$ est infime et vaut $\sim \hbar^2/(2ma_0^2)\,(a_0\Delta k)^2$, soit environ $10^{-12}\,\mathrm{eV}$.

Ces restrictions en tête, il est possible de mettre en évidence un phénomène paradoxal en apparence, tout en se souvenant que $\hbar\vec{k}$ n'est pas l'impulsion de l'électron et que, par voie de conséquence, $\hbar\vec{k}/m$ n'est pas la vitesse de ce dernier, donnée au contraire par (32.15). Ici, quand t augmente, $\vec{k}$ augmente : la variation de la vitesse avec le temps est donc dans le même sens que sa variation vis-à-vis de $\vec{k}$. Soit maintenant une bande de dispersion $\varepsilon_n(\vec{k})$ dont la variation, le long d'une direction de $\tilde{\mathcal{B}}$ a l'allure d'un cosinus inversé, minimum en centre de zone, maximum en bord de zone. Quand on part du centre, ε_n augmente avec $\vec{k}$, sa dérivée est positive, puis s'annule au bord

[12] C'est une conséquence du fait que les lois de dispersion sont périodiques dans le réseau réciproque, ou encore que la pseudo-impulsion est définie à un vecteur près $\in \tilde{\mathcal{B}}$.

de zone ; comme $\vec{k}$ et t varient dans le même sens, on peut dire que l'accélération est positive jusqu'au milieu de zone, négative ensuite. En quelque sorte, $\vec{k}$ augmentant uniformément (tout comme le temps), la vitesse augmente (l'accélération est donc positive), puis diminue (l'accélération devient négative) ; en particulier, près du bord de zone, l'accélération est dirigée en sens inverse de la force, exactement comme si l'électron avait une masse négative (cette analyse est une autre justification de la relation exprimant la masse effective de l'électron comme l'inverse de la dérivée seconde de ε_n, au facteur $\hbar^2$ près). L'explication de ce comportement paradoxal est toujours la même : la force totale agissant sur l'électron ne se réduit pas à celle du champ extérieur appliqué mais contient également le potentiel périodique de réseau, incorporé dans la loi de dispersion $\varepsilon_n(\vec{k})$. En pratique, comme l'augmentation de $\vec{k}$ est limitée par le fait que l'on doit raisonner à une échelle de temps petite par rapport au temps de collision, on se bornera à retenir qu'un électron partant près du centre de zone est accéléré par le champ, cependant qu'un autre partant tout près du bord de zone a, de fait, une masse négative, est envoyé par le champ vers d'autres états de masse négative et, au total, est freiné par l'introduction de ce dernier (l'énergie de l'électron commence par croître jusqu'au bord de zone puis, se retrouvant avec l'impulsion $-\frac{\vec{K}}{2}$, l'électron voit son énergie diminuer).

C'est d'ailleurs à ce propos que l'on peut justifier la notion de *réflexion de Bragg.* Soit un électron initialement tout près du bord de zone ($\vec{k} \lesssim \frac{\vec{K}}{2}$), donc à proximité du plan de Bragg de $\vec{K}$; sous l'effet du champ, sa pseudo-impulsion augmente et finit par dépasser $\frac{\vec{K}}{2}$, $\vec{k} = \frac{\vec{K}}{2} + \delta\vec{k}$, mais comme $\vec{k}$ et $\vec{k} - \vec{K}$ représente le même état physique, tout se passe comme si l'électron avait une pseudo-impulsion égale à $-\frac{\vec{K}}{2} + \delta\vec{k}$: la (pseudo-)vitesse s'étant inversée au voisinage du plan, il est naturel de parler de *réflexion*, comme sur un miroir[13].

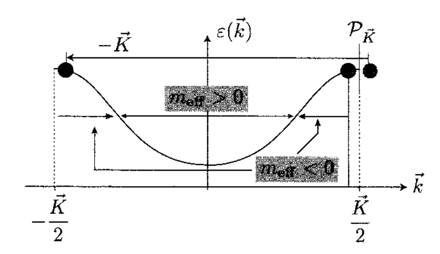


Figure 32.2: Illustration de la réflexion d'un électron sur un plan de Bragg $\mathcal{P}_{\vec{K}}$.

Le courant électrique induit par le champ extérieur se calcule à partir de la vitesse des électrons, chaque état $\vec{k}$ pouvant abriter deux électrons de spins opposés. $f(\varepsilon)$ étant

[13]Il s'agit bien seulement d'une image, puisque précisément $\hbar\vec{k}$, qui change de signe, *n'est pas* la vraie impulsion de l'électron, donc $\hbar\vec{k}/m$ n'est pas la vitesse. Celle-ci est toujours $\vec{v}_n(\vec{k})$ et s'annule continûment en t_1, si t_1 désigne l'instant où $\vec{k}(t)$ arrive en $\vec{K}/2$, puisque $\vec{\nabla}\varepsilon(\vec{K}/2) = 0$. Il reste que comme $\varepsilon(\vec{k})$ (et sa dérivée) sont des fonctions périodiques en $\vec{k}$, la (vraie) vitesse *oscille* : là où la concavité de ε est dirigée vers le bas, l'accélération est en sens contraire de la force (masse effective négative). Ceci donne lieu aux *oscillations de Bloch* (voir problème 30.7.13, où ce phénomène est transposé à des atomes situés dans un réseau optique).

la distribution de Fermi - Dirac (définie par unité de volume), on a :

$$\boxed{\vec{j} = 2\int \frac{\mathrm{d}^3k}{(2\pi)^3}\, e\, f\big(\varepsilon(\vec{k})\big)\vec{v}(\vec{k})} \tag{32.17}$$

soit à température nulle, f devenant une marche :

$$\vec{j} = 2\int_{\text{états occupés}} \frac{\mathrm{d}^3k}{(2\pi)^3}\, e\,\vec{v}(\vec{k}) \ . \tag{32.18}$$

Cette intégrale est nulle si on somme sur toute une bande :

$$\int_{\text{bande complète}} \frac{\mathrm{d}^3k}{(2\pi)^3}\, e\,\vec{v}(\vec{k}) = 0 \ , \tag{32.19}$$

puisque les deux états $\vec{k}$ et $-\vec{k}$ jouent le même rôle et que $\vec{v}$ est impair en tant que gradient d'une fonction paire. Pour une bande *partiellement* remplie, on peut donc écrire :

$$\int_{\text{états occupés}} \frac{\mathrm{d}^3k}{(2\pi)^3}\, e\,\vec{v}(\vec{k}) + \int_{\text{états vides}} \frac{\mathrm{d}^3k}{(2\pi)^3}\, e\,\vec{v}(\vec{k}) = 0 \ , \tag{32.20}$$

de sorte que le courant (32.18) est aussi égal à :

$$\vec{j} = 2\int_{\text{états vides}} \frac{\mathrm{d}^3k}{(2\pi)^3}\, (-e)\,\vec{v}(\vec{k}) \ . \tag{32.21}$$

Ce résultat d'apparence banale est important ; il montre que le courant est le même que si les états occupés étaient vides, et tous les états vides occupés par des particules de charge opposée à l'électron. Ces particules sont appelées *trous*. Ainsi, selon la commodité, on pourra raisonner soit avec les électrons d'une bande, soit avec leurs “antiparticules”.

32.3.2 Mouvement dans un champ magnétique constant

En présence d'un champ magnétique constant $\vec{\mathcal{B}}$ parallèle à Oz, les équations semi-classiques se réduisent à :

$$\dot{\vec{r}} \equiv \vec{v}_n(\vec{k}) = \frac{1}{\hbar}\frac{\partial \varepsilon_n}{\partial \vec{k}} \ , \tag{32.22}$$

$$\hbar\dot{\vec{k}} = e\,\vec{v}_n(\vec{k}) \times \vec{\mathcal{B}}(\vec{r}) \ . \tag{32.23}$$

La seconde équation montre que $\dot{\vec{k}}$ est perpendiculaire à $\vec{\mathcal{B}}$, autrement dit, la composante du vecteur $\vec{k}$ le long du champ est une constante du mouvement. Techniquement, le produit scalaire $\vec{k}.\vec{\mathcal{B}}$ a une dérivée nulle :

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\vec{k}.\vec{\mathcal{B}} = 0 \ . \tag{32.24}$$

L'énergie moyenne est aussi une constante du mouvement. En effet, la variation de $\vec{k}$ avec t entraîne *a priori* celle de $\varepsilon_n(\vec{k})$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\varepsilon_n(\vec{k}) = (\vec{\nabla}_{\vec{k}}\,\varepsilon_n)\,\frac{\mathrm{d}\vec{k}}{\mathrm{d}t}\ , \tag{32.25}$$

mais compte tenu de (32.22) et de (32.23), il vient :

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\varepsilon_n(\vec{k}) = e\hbar\vec{v}_n.(\vec{v}_n \times \vec{\mathcal{B}}) = 0\ ; \tag{32.26}$$

tout comme en Dynamique classique, la force magnétique ne travaille pas.

Le vecteur $\vec{k}$ est cependant variable dans le temps, d'où l'existence d'une orbite (trajectoire) dans l'espace $\vec{k}$, ligne décrite par l'extrémité de ce vecteur au cours du temps. L'existence des deux constantes du mouvement k_z et ε_n entraîne que ces trajectoires sont les intersections entre des plans perpendiculaires à $\vec{\mathcal{B}}$ et les surfaces d'énergie constante. Le sens de rotation sur ces orbites est fixé par $\dot{\vec{k}}$ (qui donne la vitesse sur la trajectoire dans l'espace $\vec{k}$) et donc, d'après (32.23), dépend de $\vec{v}$, lui-même déterminé par l'orientation de $\vec{\nabla}\varepsilon_n$ (qui détermine la normale à la surface d'énergie constante). Pour l'électron (de charge négative !), la rotation de $\vec{k}$ est dans le sens positif autour du champ $\vec{\mathcal{B}}$ si les états extérieurs à la surface ont une énergie inférieure à celle de ceux situés à l'intérieur.

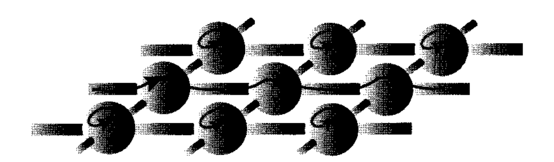


Figure 32.3: Orbites ouvertes et fermées pour deux orientations du champ magnétique, qui est normal au plan de chaque trajectoire.

Bien évidemment, ces orbites ne sont pas nécessairement fermées, d'autant plus que la topologie des surfaces (de Fermi) peut être très complexe. Par exemple, dans un schéma en zones étendues, la surface de Fermi peut ressembler à des boules vaguement sphériques (centrées sur les nœuds de $\tilde{\mathcal{B}}$) reliées par des tubes perpendiculaires aux plans de Bragg[14] (voir fig. 32.3) ; en pareil cas, suivant l'orientation du champ θ, les orbites peuvent être ouvertes ou fermées, donnant finalement une conductivité variable[15] avec

[14]Voir les ouvrages classiques pour des dessins plus explicites, et notamment [236].

[15]La dépendance de la conductivité avec le champ magnétique a été découverte par William Thomson (Lord Kelvin) en 1857. L'effet sommairement décrit ci-dessus doit être distingué de la *magnétorésistance géante*, découverte par Fert [253] et Grünberg [254] en 1988, qui désigne la chute très rapide de la résistance dans des composés formés de multicouches ferromagnétiques séparées par une mince couche (quelques angströms) non magnétique. Pour ces composés, la résistance $R(\mathcal{B})$ décroît rapidement en fonction du seul *module* du champ magnétique appliqué, une fois l'orientation de celui-ci judicieusement choisie : le rapport $R(\mathcal{B})/R(0)$ passe de 1 à $\gtrsim 0,5$ pour un champ de l'ordre de 20 kGauss [253].

Pour cette découverte aux applications multiples (disques durs !), Albert Fert et Peter Grünberg ont reçu le prix Nobel 2007.

le champ (magnétorésistance).

Il est également possible de trouver la trajectoire dans l'espace physique. La variation de $\vec{r}_\perp \stackrel{\text{déf}}{=} \vec{r} - \frac{\vec{\mathcal{B}}.\vec{r}}{\mathcal{B}} \equiv \vec{r} - \vec{r}_\parallel$, qui donne la projection du mouvement dans un plan perpendiculaire au champ, s'obtient comme suit ; le deuxième membre de (32.23) devient :

$$\hbar\dot{\vec{k}} = e\,\dot{\vec{r}}_\perp \times \vec{\mathcal{B}} \ , \tag{32.27}$$

et une simple intégration en temps de (32.23) donne :

$$\frac{\hbar}{e}\,[\vec{k}(t) - \vec{k}(0)] = [\vec{r}_\perp(t) - \vec{r}_\perp(0)] \times \vec{\mathcal{B}} \ . \tag{32.28}$$

La solution de cette équation est[16] :

$$\boxed{\vec{r}_\perp(t) - \vec{r}_\perp(0) = \frac{\hbar}{e\mathcal{B}^2}\,[\vec{k}(t) - \vec{k}(0)] \times \vec{\mathcal{B}}} \tag{32.30}$$

Il en résulte que la projection des orbites "spatiales" sur un plan perpendiculaire au champ s'obtient en tournant de $\frac{\pi}{2}$ autour du champ les orbites pour $\vec{k}$ et en les multipliant par le facteur dimensionnel $\frac{\hbar}{e\mathcal{B}^2}$. Dans le cas des électrons libres, les surfaces d'énergie constante sont des sphères dont l'intersection avec les plans perpendiculaires au champ sont des cercles ; en tournant ces cercles de $\frac{\pi}{2}$, on obtient encore des cercles, en conformité avec le résultat classique. En présence du potentiel périodique, les orbites peuvent être très compliquées, en résultat de la topologie complexe des surfaces d'énergie constante.

L'obtention de la variation de la troisième coordonnée est moins simple, puisqu'il faut résoudre :

$$z(t) - z(0) = \int_0^t v_z\big(\vec{k}(t')\big)\,\mathrm{d}t' \ , \qquad v_z\big(\vec{k}(t)\big) = \hbar^{-1}\,\frac{\partial\varepsilon_n}{\partial k_z}\big(\vec{k}(t)\big) \tag{32.31}$$

Si l'électron était libre ($\varepsilon(\vec{k}) = \frac{\hbar^2\vec{k}^2}{2m}$), on aurait simplement $v_z\big(\vec{k}(t)\big) = \frac{\hbar k_z(t)}{m}$, et comme k_z est, en toute hypothèse, une constante du mouvement, on trouverait immédiatement :

$$z(t) - z(0) = \frac{\hbar k_z}{m}\,t \ , \tag{32.32}$$

reproduisant le mouvement uniforme classique le long du champ. En présence du potentiel périodique, la symétrie sphérique est perdue et $\frac{\partial\varepsilon_n}{\partial k_z}$ contient en général les trois composantes k_x, k_y et k_z.

---

[16] Soit l'équation :

$$\vec{A} \times \vec{B} = \vec{C} \ . \tag{32.29}$$

En posant $\vec{A} = \vec{A}_\perp + \vec{A}_\parallel$, (32.29) devient $\vec{A}_\perp \times \vec{B} = \vec{C}$. En construisant la figure illustrant cette relation vectorielle, on voit immédiatement que $\vec{A}_\perp$, $\vec{B}$ et $\vec{C}$ forment un trièdre trirectangle et que $\vec{A}_\perp$ est égal à $\frac{\vec{B}\times\vec{C}}{\|\vec{B}\|^2}$.

32.4 Conductivité statique d'un métal

Le modèle classique de Drude met bien en évidence un fait évident physiquement : si les particules responsables du transport de courant électrique n'étaient pas freinées, il ne pourrait y avoir de loi d'Ohm, puisque sous l'effet du champ électrique, leur vitesse augmenterait linéairement en temps. L'existence d'un courant constant dans un conducteur en régime permanent soumis à un champ électrique constant exige que la vitesse des porteurs – en tout cas celle qui produit le courant – soit elle-même constante.

Le modèle semi-classique sommairement exposé ci-dessus prétend seulement donner la dynamique des électrons dans un cristal entre deux collisions, quelle que soit la nature de ces dernières (défauts structuraux et/ou chimiques, phonons, etc.). Il permet d'obtenir des équations dynamiques où $\varepsilon(\vec{k})$ joue le rôle d'un Hamiltonien au sens classique et où $\hbar\vec{k}$ a toutes les apparences d'un moment conjugué (impulsion). Ces analogies permettent de construire une théorie simple du transport électrique, en associant les équations dynamiques semi-classiques précédentes à des termes représentant des collisions au sens large et selon une méthode universelle, suivie notamment en Mécanique statistique classique et adaptée au cadre quantique remodelé par le modèle semi-classique.

La description mécanique précise des collisions passe par un traitement *statistique* reposant sur des hypothèses plausibles incorporant de façon effective les degrés de liberté laissés en coulisse. La quantité centrale est alors la fonction de distribution ρ donnant, à un certain instant t la (densité de) probabilité de trouver un système avec des valeurs déterminées pour les grandeurs dynamiques précisant son état. Ici, le mouvement de chaque électron "semi-classique" est défini par les deux vecteurs $\vec{r}$ et $\vec{k}$, dont la variation dans le temps définit l'*orbite* de l'électron dans cet espace des phases ; la fonction de distribution est donc notée plus précisément $\rho(\vec{r}, \vec{k}, t)$. Le nombre $\mathrm{d}N$ d'électrons ayant à l'instant t la position $\vec{r}$ à d^3r près et le vecteur d'onde $\vec{k}$ à d^3k près est :

$$\mathrm{d}N = \rho(\vec{r}, \vec{k}, t)\, 2\, \frac{\mathrm{d}^3r\, \mathrm{d}^3k}{(2\pi)^3} \ , \tag{32.33}$$

où le facteur 2 prend en compte le spin. À l'équilibre thermodynamique, pour les fermions que sont les électrons, ρ n'est autre que la distribution de Fermi - Dirac donnée en (32.12).

En l'absence de collisions, $\rho(\vec{r}, \vec{k}, t)$ obéit à une équation purement mécanique (l'un des avatars de l'équation de Liouville) que l'on écrit simplement sur la base d'un argument de conservation. En effet, l'intégrale de ρ dans tout l'espace des phases est une constante dans le temps, égale au nombre total d'électrons :

$$\int_{\mathbb{R}^3} \mathrm{d}^3r \int_{\mathbb{R}^3} \frac{\mathrm{d}^3k}{4\pi^3}\, \rho(\vec{r}, \vec{k}, t) = N \ . \tag{32.34}$$

Dès lors, le mouvement des N électrons (dans l'espace des phases) est analogue à celui d'un fluide sans perte pour lequel il existe une équation locale de conservation :

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}.\vec{\mathcal{J}} = 0 \ . \tag{32.35}$$

$\vec{\mathcal{J}}$ désigne le courant dans l'espace des phases ; c'est le produit de la densité ρ par la vitesse de déplacement d'un point de cet espace, défini par les six composantes des vecteurs $\vec{r}$ et $\vec{k}$ – celles de la vitesse correspondante étant les six composantes de $\dot{\vec{r}}$ et de $\dot{\vec{k}}$. Quant à la divergence, c'est le gradient à six composantes, réunion de $\vec{\nabla}_{\vec{r}}$ et de $\vec{\nabla}_{\vec{k}}$. Dès lors, (32.35) s'écrit :

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}_{\vec{r}}\,(\rho\dot{\vec{r}}) + \vec{\nabla}_{\vec{k}}\,(\rho\dot{\vec{k}}\,) = 0 \ . \tag{32.36}$$

Chaque gradient donne un premier terme proportionnel à ρ, $\rho\,(\vec{\nabla}_{\vec{r}}\,\dot{\vec{r}} + \vec{\nabla}_{\vec{k}}\,\dot{\vec{k}})$; s'il existait un authentique Hamiltonien déterminant la dynamique par les équations canoniques, la parenthèse serait globalement nulle : le premier terme serait de la forme $\partial_q(\partial_p H)$ et le second serait $\partial_p(-\partial_q H)$. Ici les choses se présentent différemment. En fait, le premier terme de la parenthèse est nul (ε ne dépend pas de $\vec{r}$) :

$$\vec{\nabla}_{\vec{r}}\,\dot{\vec{r}} \equiv \vec{\nabla}_{\vec{r}}\,\vec{v} = \hbar^{-1}\,\vec{\nabla}_{\vec{r}}\,\varepsilon(\vec{k}) \equiv 0 \ ; \tag{32.37}$$

quant au second terme, il vaut :

$$e\hbar^{-1}\vec{\nabla}_{\vec{k}}\,(\vec{\mathcal{E}} + \vec{v} \times \vec{\mathcal{B}}) = e\hbar^{-1}\vec{\nabla}_{\vec{k}}\,(\vec{\nabla}_{\vec{k}}\,\varepsilon \times \vec{\mathcal{B}}) \ . \tag{32.38}$$

C'est un produit mixte impliquant deux fois le même vecteur ($\vec{\nabla}_{\vec{k}}$), qui est donc nul.

En définitive, l'équation décrivant l'écoulement du fluide décrit par les variables dynamiques *ad hoc* $\dot{\vec{r}} \equiv \vec{v}$ et $\dot{\vec{k}}$ obéissant aux équations du mouvement semi-classiques (32.10) et (32.11) est, en l'absence de collisions :

$$\boxed{\frac{\partial \rho}{\partial t} + \dot{\vec{r}}.\vec{\nabla}_{\vec{r}}\rho + \dot{\vec{k}}.\vec{\nabla}_{\vec{k}}\rho = 0} \tag{32.39}$$

qui est une équation de Liouville *effective*.

Les collisions modifient position et "impulsion" et induisent des pertes au sens où un électron qui devrait se trouver en ($\vec{r}$, $\vec{k}$) à l'instant t s'il suivait les équations strictement dynamiques, ne s'y retrouve pas (mais, bien évidemment, se trouve ailleurs), et inversement : on retrouve en ($\vec{r}$, $\vec{k}$) des électrons qui, sans collisions, seraient en d'autres points de l'espace des phases. Il en résulte que, en présence des collisions, le second membre de (32.36) n'est pas nul, et est précisément égal à la variation temporelle de ρ sous l'effet des collisions, ce que l'on traduit formellement par :

$$\frac{\partial \rho}{\partial t} + \dot{\vec{r}}.\vec{\nabla}_{\vec{r}}\rho + \dot{\vec{k}}.\vec{\nabla}_{\vec{k}}\rho = \left(\frac{\partial \rho}{\partial t}\right)_{\text{coll}} \ . \tag{32.40}$$

Ceci n'avance guère, tant que l'on ne connaît pas le second membre. Une hypothèse raisonnable consiste à dire que cette variation sous l'effet des collisions est d'autant plus grande que ρ diffère beaucoup de la distribution d'équilibre ρ_{m} du métal situé dans le même environnement thermodynamique (en particulier, même température, éventuellement définie dans le cadre d'un équilibre local). On est donc enclin à écrire :

$$\left(\frac{\partial \rho}{\partial t}\right)_{\text{coll}} \propto (\rho - \rho_{\text{m}}) \ , \tag{32.41}$$

la constante de proportionnalité étant l'inverse d'un temps ; enfin, pour qu'il s'agisse d'une relaxation, il convient d'introduire le bon signe, conduisant à la forme acceptable :

$$\left(\frac{\partial\rho}{\partial t}\right)_{\text{coll}} = -\frac{1}{\tau}\,(\rho - \rho_{\text{m}}) \ , \tag{32.42}$$

où τ, positif, est appelé temps de relaxation. *A priori*, c'est une fonction de $\vec{r}$ (si le milieu est inhomogène) et de $\vec{k}$ (dépendance en énergie du temps de relaxation). En définitive, dans cette description, dite *du temps de relaxation*, l'équation d'évolution de ρ est la suivante :

$$\boxed{\frac{\partial\rho}{\partial t} + \dot{\vec{r}}.\vec{\nabla}_{\vec{r}}\rho + \dot{\vec{k}}.\vec{\nabla}_{\vec{k}}\rho = -\frac{1}{\tau}\,(\rho - \rho_{\text{m}})} \tag{32.43}$$

Avant d'exploiter une telle forme simplifiée, il est utile de donner quelques arguments – très insuffisants, comme on va le voir – permettant d'en conforter la légitimité [231]. Comme on l'a dit, ce sont les collisions au sens large qui viennent perturber le bilan nul résultant d'une dynamique strictement mécanique. Introduisons la probabilité par unité de temps, w, pour qu'un électron subisse une collision entre deux instants t et $t+\mathrm{d}t$, et la probabilité $P(t)$ qu'il n'ait subi aucune collision pendant un intervalle de temps t. En écrivant ce qui peut se passer entre t et $t+\mathrm{d}t$, on arrive à $P(t+\mathrm{d}t) = (1-w\mathrm{d}t)P(t)$, dont la solution est $P(t) = \mathrm{e}^{-wt}$; le temps moyen entre deux collisions est $\tau_{\text{coll}} = \int_{\mathbb{R}_+} P(t)\mathrm{d}t$, soit w^{-1}.

La question est de trouver les électrons semi-classiques qui se retrouvent en $\vec{r}$ et $\vec{k}$ à l'instant t, sachant qu'ils étaient en $\vec{r}_0$ et $\vec{k}_0$ à un instant antérieur t_0. En raison de la dispersion des trajectoires dues aux collisions, le bilan dans le petit volume autour de $(\vec{r}, \vec{k})$ contient un terme de peuplement (les électrons qui, en l'absence de collisions, seraient ailleurs) et un terme de dépeuplement (les électrons qui devraient s'y trouver mais qui, à cause des collisions, sont partis ailleurs). La probabilité introduite ci-dessus permet seulement de calculer ce deuxième terme – et c'est pourquoi, entre autres raisons, la justification de (32.43) sera loin d'être complète.

Faisons maintenant l'hypothèse d'un équilibre *local*, en raisonnant sur un domaine de taille mésoscopique[17] pour lequel cela a un sens de définir les grandeurs de nature statistique (entropie, température,...), et en admettant que ce petit domaine est effectivement à l'équilibre à tout instant ; dans ces conditions, la fonction de densité en $(\vec{r}_0, \vec{k}_0)$ à l'instant t_0 est la fonction ρ_{m} satisfaisant l'équation de Liouville en l'absence de collisions (32.39). Soit maintenant un électron ayant subi une collision entre $t_0 - \mathrm{d}t'$ et n'ayant subi aucune collision entre t_0 et t ; posant $t' = t - t_0$, la probabilité correspondante est $w\mathrm{d}t'\,\mathrm{e}^{-wt'}$. Les électrons arrivant en $(\vec{r}, \vec{k})$ à l'instant t s'obtiennent en sommant sur tous les instants antérieurs la fonction de l'équilibre local, affectée de ce

[17] Il s'agit à nouveau d'une opération de moyenne à gros grain (*coarse graining*) qui, ici, est bien compatible avec les échelles de longueurs définies pour arriver au modèle semi-classique.

poids (probabilité), d'où[18] :

$$\rho(\vec{r},\,\vec{k},\,t) = \int_0^{+\infty} \rho_{\rm m}(\vec{r}_0,\,\vec{k}_0,\,t-t')\,w{\rm d}t'\,{\rm e}^{-wt'} \quad . \tag{32.44}$$

Écrivant cette équation en $t + \delta t$, retranchant membre à membre, divisant par δt et prenant la limite $\delta t = 0$, on obtient au premier membre la dérivée totale de $\rho(\vec{r},\,\vec{k},\,t)$, c'est-à-dire strictement le premier membre de (32.39). Après une intégration par parties, le second membre devient $-w[\rho(\vec{r},\,\vec{k},\,t) - \rho_{\rm m}(\vec{r},\,\vec{k},\,t)]$; comme $w = 1/\tau_{\rm coll}$, l'équation (32.43) est retrouvée, à condition d'identifier le temps de relaxation τ qui y figure et le temps moyen entre deux collisions, $\tau_{\rm coll}$.

Après cette brève (et partielle) justification de l'approximation du temps de relaxation, l'équation (32.43) va maintenant être utilisée pour obtenir la conductivité électrique statique d'un métal, mais avant de continuer, un commentaire s'impose. Telle qu'elle est écrite, l'équation (32.43) est très classique [231] et, *mutatis mutandis*, est utilisée pour décrire des phénomènes, *classiques* eux aussi, extrêmement variés (relaxation après un choc thermique, dilution d'une inhomogénéité de densité dans un fluide,...). Dit autrement, l'équation peut laisser penser que la Mécanique quantique est bien loin. Évidemment, il n'en est rien : l'aspect quantique est bel et bien présent, d'une part parce que les équations (32.10) et (32.11) impliquent fondamentalement la grandeur quantique essentielle, à savoir la relation de dispersion $\varepsilon(\vec{k}\,)$ d'un électron dans le potentiel périodique – dont on a vu le rôle crucial pour l'explication de la distinction isolant/conducteur –, d'autre part en vertu du fait que c'est la fonction de Fermi - Dirac qui figure en tant que densité d'équilibre au second membre.

Pour calculer sa conductivité σ, on suppose que le solide est à une température donnée homogène (pas de gradient de température) et que le champ électrique appliqué est lui aussi homogène (pas de gradient de champ électrique) ; dans ces conditions, $\rho_{\rm m}$ est la distribution de Fermi - Dirac, f. Enfin, on s'intéresse uniquement au régime permanent, après extinction de tous les transitoires. Ainsi, dans un conducteur homogène (ρ indépendant de $\vec{r}$), le régime permanent est décrit une certaine fonction $\rho_{\rm p}(\vec{k}\,)$ elle-même déterminée par (32.43) avec $\frac{\partial\rho_{\rm p}}{\partial t} = 0$ et $\vec{\nabla}_{\vec{r}}\rho_{\rm p} = 0$:

$$\dot{\vec{k}}.\vec{\nabla}_{\vec{k}}\rho_{\rm p} = -\frac{1}{\tau}\,(\rho_{\rm p} - f) \quad \Longleftrightarrow \quad \frac{e}{\hbar}\,\vec{\mathcal{E}}.\vec{\nabla}_{\vec{k}}\rho_{\rm p} = -\frac{1}{\tau}\,(\rho_{\rm p} - f) \ , \tag{32.45}$$

où (32.11) a été utilisé pour arriver à la deuxième égalité, qui peut se récrire :

$$\boxed{\rho_{\rm p}(\vec{k}\,) = f - \frac{e\tau}{\hbar}\,\vec{\mathcal{E}}.\vec{\nabla}_{\vec{k}}\,\rho_{\rm p}(\vec{k}\,)} \tag{32.46}$$

La pseudo-impulsion $\vec{k}$ est présente dans f à droite par l'intermédiaire de la loi de dispersion $\varepsilon(\vec{k}\,)$. Sous cette forme, on voit qu'il existe *a priori* un développement en puissances

[18]Remarquer que si w est arbitrairement grand (collisions ultra-rapides), le poids est une fonction de Dirac, de sorte que (32.44) donne $\rho(\vec{r},\,\vec{k},\,t) = \rho_{\rm m}(\vec{r},\,\vec{k},\,t)$, montrant que l'équilibre se rétablit alors instantanément.

du champ, que l'on peut formellement écrire en itérant :

$$\rho_{\mathrm{P}}(\vec{k}) = f - \frac{e\tau}{\hbar}\,\vec{\mathcal{E}}.\vec{\nabla}_{\vec{k}}f + \big(\frac{e\tau}{\hbar}\big)^2\,(\vec{\mathcal{E}}.\vec{\nabla}_{\vec{k}})^2\,f + ... \tag{32.47}$$

S'en tenant désormais à l'ordre le plus bas en champ, il vient :

$$\boxed{\rho_{\mathrm{P}}(\vec{k}) \simeq f - \frac{e\tau}{\hbar}\,\vec{\mathcal{E}}.\vec{\nabla}_{\vec{k}}\,f} \tag{32.48}$$

expression qui donne la réponse *linéaire* du métal à l'application du champ et qui, dans le but de calculer la conductivité, est méthodologiquement nécessaire et suffisant : écrire $\vec{j} = \sigma\vec{\mathcal{E}}$, c'est *déjà* se placer dans un cadre de réponse linéaire.

Afin de rendre les résultats plus transparents, on utilise :

$$\vec{\nabla}_{\vec{k}}f = \left(\frac{\partial f}{\partial\varepsilon}\right)\,\vec{\nabla}_{\vec{k}}\,\varepsilon \,=\, \hbar\vec{v}\,\frac{\partial f}{\partial\varepsilon}\ , \tag{32.49}$$

de sorte que (32.48) devient :

$$\rho(\vec{k}) = f - e\tau\,\vec{\mathcal{E}}.\vec{v}\,\frac{\partial f}{\partial\varepsilon}\ . \tag{32.50}$$

Il est bien entendu que ε désigne partout la fonction $\varepsilon(\vec{k})$ et que la vitesse dépend de $\vec{k}$. Maintenant, il suffit de revenir à l'expression du courant de charge en régime permanent (compté par unité de volume[19]) :

$$\vec{j} = e\int\frac{\mathrm{d}^3k}{4\pi^3}\,\rho_{\mathrm{P}}(\vec{k})\vec{v}(\vec{k}) \tag{32.51}$$

pour obtenir :

$$\vec{j} = e\int\frac{\mathrm{d}^3k}{4\pi^3}\Big(f - e\tau\vec{\mathcal{E}}.\vec{v}\frac{\partial f}{\partial\varepsilon}\Big)\vec{v}\ . \tag{32.52}$$

L'intégration avec f donne évidemment zéro (pas de courant à l'équilibre en l'absence de champ). En explicitant les produits scalaires à l'aide des coordonnées cartésiennes, il vient :

$$j_\alpha = -e^2\sum_{\beta=x,\,y,\,z}\int\frac{\mathrm{d}^3k}{4\pi^3}\,\tau v_\alpha\mathcal{E}_\beta v_\beta\frac{\partial f}{\partial\varepsilon} \equiv \sum_\beta\sigma_{\alpha\,\beta}\mathcal{E}_\beta\ , \tag{32.53}$$

où le tenseur de conductivité $\sigma_{\alpha\,\beta}$ se trouve défini par la dernière égalité, généralisant la relation usuelle et incorporant une éventuelle anisotropie de conductivité[20] :

$$\boxed{\vec{j} = \hat{\sigma}\,\vec{\mathcal{E}} \qquad \hat{\sigma} = -e^2\int\frac{\mathrm{d}^3k}{4\pi^3}\,\tau\big(\varepsilon(\vec{k})\big)\vec{v}\otimes\vec{v}\,\frac{\partial f}{\partial\varepsilon}} \tag{32.54}$$

[19]Comme f est ici indépendant de $\vec{r}$, l'intégration sur $\vec{r}$ donne un simple facteur extensif, le volume de l'échantillon.

[20]Noter que σ a le bon signe puisque $\frac{\partial f}{\partial\varepsilon}$ est négatif.

Dans un cristal cubique, prenant x, y et z le long des arêtes du cube, les trois éléments diagonaux $\sigma_{\alpha\alpha}$ sont égaux entre eux : le tenseur de conductivité est donc un simple scalaire, traduisant l'isotropie (réalisée aussi bien sûr dans le cas libre).

Il est possible de faire apparaître le tenseur de masse effective introduit dans le chapitre 30 (sous-section 30.4.1, éq. (30.132)). On note d'abord que, aux températures usuelles, le rapport $k_{\text{B}}T/\varepsilon_{\text{F}}$ est très petit devant 1, de sorte que la fonction f est très voisine d'une marche. Il en résulte que $\frac{\partial f}{\partial \varepsilon} \simeq -\delta(\varepsilon - \varepsilon_{\text{F}})$, ce qui autorise à sortir $\tau(\varepsilon)$ de l'intégrale et donne :

$$\hat{\sigma} \simeq -e^2\tau(\varepsilon_{\text{F}}) \int \frac{\mathrm{d}^3k}{4\pi^3}\, \vec{v} \otimes \vec{v}\frac{\partial f}{\partial \varepsilon} \ . \tag{32.55}$$

Par ailleurs, avec $\vec{v}\,\frac{\partial f}{\partial \varepsilon} = \hbar^{-1}\vec{\nabla}_{\vec{k}} f\big(\varepsilon(\vec{k})\big)$, l'élément $\alpha\beta$ de $\hat{\sigma}$ est :

$$\sigma_{\alpha\beta} = -\frac{e^2}{\hbar}\tau(\varepsilon_{\text{F}}) \int \frac{\mathrm{d}^3k}{4\pi^3}\, v_\alpha \frac{\partial f\big(\varepsilon(\vec{k})\big)}{\partial k_\beta} \tag{32.56}$$

et une intégration par parties donne :

$$\sigma_{\alpha\beta} = \frac{e^2}{\hbar}\tau(\varepsilon_{\text{F}}) \int \frac{\mathrm{d}^3k}{4\pi^3}\, f\big(\varepsilon(\vec{k})\big)\frac{\partial v_\alpha}{\partial k_\beta} \ . \tag{32.57}$$

Le tenseur de masse effective d'éléments $m_{\alpha\beta}$ est défini comme (voir (30.132)) :

$$(m^{-1})_{\alpha\beta} = \hbar^{-2}\,\frac{\partial^2 \varepsilon}{\partial k_\alpha \partial k_\beta} = \hbar^{-1}\,\frac{\partial v_\alpha}{\partial k_\beta} = \frac{\partial^2 \varepsilon}{\partial k_\beta \partial k_\alpha} = \hbar^{-1}\,\frac{\partial v_\beta}{\partial k_\alpha} \ . \tag{32.58}$$

d'où une autre expression de $\sigma_{\alpha\beta}$:

$$\sigma_{\alpha\beta} = e^2\tau(\varepsilon_{\text{F}}) \int \frac{\mathrm{d}^3k}{4\pi^3}\, f\big(\varepsilon(\vec{k})\big)\,(m^{-1})_{\alpha\beta} \ . \tag{32.59}$$

σ, tout comme m, est un tenseur symétrique. À très basse température, on peut enfin remplacer f par la fonction de Heaviside $\Theta(\varepsilon_{\text{F}}-\varepsilon)$, donnant les éléments de la conductivité sous la forme :

$$\boxed{\sigma_{\alpha\beta} = e^2\tau(\varepsilon_{\text{F}}) \int_{\text{états occupés}} \frac{\mathrm{d}^3k}{4\pi^3}\,(m^{-1})_{\alpha\beta}} \tag{32.60}$$

Comme $(m^{-1})_{\alpha\beta}$ est la dérivée d'une fonction périodique, sa moyenne sur une maille primitive est nulle ; (32.60) s'écrit donc tout autant :

$$\sigma_{\alpha\beta} = e^2\tau(\varepsilon_{\text{F}}) \int_{\text{états vides}} \frac{\mathrm{d}^3k}{4\pi^3}\,(-m^{-1})_{\alpha\beta} \ . \tag{32.61}$$

Dans cette optique, ce sont les trous positifs qui conduisent le courant et ont une masse effective changée de signe par rapport à celle des électrons. Par ailleurs, si $(m^{-1})_{\alpha\beta}$ est scalaire et indépendant de $\vec{k}$:

$$(m^{-1})_{\alpha\beta} = \frac{1}{m_{\text{eff}}}\delta_{\alpha\beta} \ , \tag{32.62}$$

alors :

$$\sigma_{\alpha\beta} = \sigma\,\delta_{\alpha\beta}\ , \qquad \sigma = e^2\tau(\varepsilon_{\mathrm{F}})\frac{1}{m_{\mathrm{eff}}}\int_{\text{états occupés}} \frac{\mathrm{d}^3k}{4\pi^3} \tag{32.63}$$

l'intégrale n'est autre que le nombre d'électrons par unité de volume, soit la densité n_{e}. En définitive, on obtient :

$$\boxed{\sigma = \frac{n_{\mathrm{e}}\,e^2\tau}{m_{\mathrm{eff}}}} \tag{32.64}$$

Ceci n'est autre que le résultat de Drude, (32.4), corrigé de la masse effective. Cette modification d'apparence anodine est en fait de première importance : elle le point de départ des recherches en physique des matériaux destinées à réduire la masse effective autant que faire se peut, afin d'obtenir de grandes valeurs pour la conductivité.

♦ *Remarques*

1. Cette théorie élémentaire de la conductivité électrique ne vaut que pour des systèmes de taille macroscopique. Le modèle semi-classique repose sur la possibilité de définir des porteurs ayant une position et une vitesse déterminées au sens défini alors (voir p. 1523), ce qui énonce clairement que l'on considère exclusivement des échelles de longueur très grandes par rapport à une distance atomique.

   On sait de nos jours réaliser des circuits de taille *nanoscopique*, pour lesquels la théorie du transport doit donc être entièrement reconsidérée (pour une discussion générale, voir [255]).

2. Fondamentalement, on s'est appuyé sur les résultats principaux de la théorie des bandes, théorie à *une* particule, dont les limites ont été brièvement discutées dans le chapitre 30, Remarque 3, p. 1411. Une autre manifestation spectaculaire des possibles effets à N-corps est l'apparition de la supraconductivité, introduite dans la section suivante. ♦

32.5 Introduction à la supraconductivité

32.5.1 Manifestations expérimentales de l'état supraconducteur

La manifestation la plus spectaculaire de la supraconductivité (*superconductivity*) est évidemment la disparition de toute résistance électrique[21] de certains métaux (mercure, étain, plomb,...) au-dessous d'une certaine température critique T_{c}, variable d'un métal à l'autre mais toujours très basse (quelques K !), au moins pour les supraconducteurs

[21] Cette découverte extraordinaire [256] a été faite en 1911 par Gilles Holst, étudiant en thèse sous la direction de Kamerlingh Onnes, sur un barreau de mercure ($T_{\mathrm{c}} = 4,15$ K).

Heike Kamerlingh Onnes (1853-1926) a reçu le Nobel en 1913 *"for his investigations on the properties of matter at low temperatures which led, inter alia, to the production of liquid helium"*.

"historiques" [22]. Une fois établi dans un tel métal, le courant peut persister sans aucune perte mesurable pendant des années, et on estime typiquement à plusieurs milliers d'années la borne *inférieure* du temps d'amortissement. C'est pourquoi la *conductibilité électrique parfaite* est considérée comme la signature de la supraconductivité.

D'autres effets sont également caractéristiques de cet état remarquable, notamment l'effet Meissner [261] où, abaissant la température d'un métal normal soumis à un champ magnétique, on voit brusquement le champ expulsé[23] du matériau au moment où il devient supraconducteur au passage de T_c (diamagnétisme parfait). Par ailleurs, le très important effet *isotopique* désigne la dépendance de la température critique par rapport à la masse ($T_c \propto M^{-1/2}$) de différents isotopes d'un même élément, découvert par E. Maxwell en 1950 [262].

Enfin, la chaleur spécifique c_V varie *exponentiellement* avec la température [263] quand $T < T_c$: $c_V \propto \mathrm{e}^{-aT_c/T}$. Comme discuté à plusieurs reprises à propos des magnons (chapitre 30, sous-section 30.6.1) ou des phonons (chapitre 31, section 31.6), ce comportement de c_V signe l'existence d'un *gap* fini d'énergie entre l'état fondamental et les premières excitations de basse énergie.

32.5.2 La théorie BCS

Il s'agit ici de donner un bref aperçu de la théorie microscopique standard de la supraconductivité [265], dite[24] *théorie BCS*, publiée en 1957, qui repose fondamentalement sur la notion de *paires de Cooper*.

En 1956, Cooper a montré [264] que la mer de Fermi est *instable* vis-à-vis de la formation de paires liées d'électrons pourvu qu'il existe une interaction attractive,

[22] L'année 1986 a connue une certaine effervescence en Physique – le moins que l'on puisse dire –, après l'annonce par Bednorz et Müller [257] de la découverte de nouveaux matériaux supraconducteurs à *haute* température critique (quelques dizaines de K). La course aux records qui s'en est suivie, faisant parfois un peu vite la *une* du *New York Times*, n'a guère favorisé la compréhension en profondeur de la physique de ces matériaux, et a provoqué une compétition dont la vivacité n'a eu d'égal que les violentes controverses qu'elle a suscitées. À l'heure actuelle, il semble qu'un consensus s'est dégagé pour admettre qu'il ne s'agit pas d'une *nouvelle* supraconductivité, et que les théories conventionnelles pourraient s'appliquer dans leurs principes à ces matériaux exotiques, tout en n'étant pas à même d'incorporer, en pratique, les spécificités de ces systèmes (forte anisotropie, densité d'états très particulière, structure cristalline, etc.) – pour différentes approches des supras-haute T_c, voir [258], [259] et [260]. Actuellement, on connaît des *supras* dont la température critique est de l'ordre de 150 K, très au-dessus de l'azote liquide (77 K).

Pour leur découverte de 1986, Bednorz et Müller ont reçu le Nobel l'année suivante.

[23] Ceci est vrai à l'échelle d'un échantillon macroscopique : en réalité, le champ pénètre sur une distance finie λ, de l'ordre de 500 Å (profondeur de pénétration, appelée *longueur de London*).

[24] L'acronyme est constitué des initiales de ses trois auteurs : Bardeen, Cooper et Schrieffer. Leon Nathan Cooper est né en 1930, John Robert Schrieffer en 1931. John Bardeen (1908-1991) est la seule personne à avoir été deux fois lauréat Nobel en Physique : en 1956 avec William Shockley et Walter Brattain pour l'invention du transistor, une deuxième fois en 1972 pour la théorie BCS.

À ce sujet, il convient de mentionner d'autres mécanismes, donnant lieu à la supraconductivité dite *non conventionnelle* (par exemple, *triplet superconductivity*, imaginé par Ginzburg en 1957) ; le premier exemple connu de matériau relevant de ce schéma est le composé organique appelé $(TMTSF)_2PF_6$, dont Jérome et al. [266] ont découvert en 1979 les propriétés fort inusuelles.

aussi petite soit-elle, entre ces deux électrons. Comme on va le voir, ce phénomène assez extraordinaire est une conséquence de l'existence du fluide constitué par les électrons de la mer de Fermi, où baignent les deux électrons candidats à une paire liée : indépendamment de l'origine mystérieuse de l'interaction *attractive* (entre deux électrons !), le fait est que, à trois dimensions d'espace, des états liés ne peuvent exister que si l'attraction est assez forte, voir par exemple le cas du puits carré sphérique[25] (chapitre 19, sous-section 19.1.4).

Afin de mettre en évidence ce résultat capital aussi simplement que possible [267], considérons deux électrons noyés dans la mer de Fermi à température nulle, ces deux électrons étant supposés interagir entre eux de manière *attractive* – sans questionner pour l'instant l'origine de cette interaction –, mais ne sentant la présence des autres électrons (ceux de la mer) que via le Principe de Pauli. Tout comme on l'a fait pour les magnons (chapitre 30, sous-section 30.6.2, notamment éq. (30.227)), on écrit la fonction d'onde de cette paire en combinaison linéaire du produit de deux ondes planes de vecteurs d'onde $\vec{k}_1$ et $\vec{k}_2$, en partant de l'idée que l'état fondamental de cette paire correspond à une impulsion totale nulle ($\vec{K} \stackrel{\text{déf}}{=} \vec{k}_1 + \vec{k}_2 = 0$, $\vec{q} \stackrel{\text{déf}}{=} \frac{1}{2}(\vec{k}_1 - \vec{k}_2)$), voir (30.226)) ; notant ainsi $\vec{k}_1 = -\vec{k}_2 = \vec{k}$, d'où $\vec{q} = \vec{k}$, le point de départ pour écrire la fonction d'onde est :

$$\sum_{\vec{k}} \Phi_{\vec{0}}(\vec{k}) \mathrm{e}^{\mathrm{i}\vec{k}.\vec{r}_1} \mathrm{e}^{-\mathrm{i}\vec{k}.\vec{r}_2} \ . \tag{32.65}$$

Telle quelle, cette expression ne convient pas, puisqu'elle ne satisfait pas l'obligation imposée par le postulat d'antisymétrisation. Avec deux électrons, on peut faire un état triplet, dont la fonction d'espace est antisymétrique, ou un état singulet, à fonction d'espace symétrique. On sait (voir l'atome d'hélium, chapitre 27, section 27.7) que quand l'interaction est répulsive, l'état triplet est toujours plus stable (d'énergie inférieure) que l'état singulet, en raison du trou de Fermi. Ici, la question posée n'a de sens que si l'interaction est attractive, une situation dans laquelle l'ordre des énergies est évidemment inversé : s'agissant de trouver l'état fondamental de la paire, il faut donc choisir un état *singulet*. Cela étant, la partie d'espace d'une bonne fonction d'onde s'écrit :

$$\boxed{\psi_0(\vec{r}_1, \vec{r}_2) = \sum_{\vec{k}, \|\vec{k}\| > k_\mathrm{F}} \Phi_{\vec{0}}(\vec{k}) \cos\left(\vec{k}.(\vec{r}_1 - \vec{r}_2)\right)} \tag{32.66}$$

à multiplier, le cas échéant, par la fonction de spin singulet $X_{0,0}$ (voir (27.54)). Dans l'expression (32.66), $\vec{k}_\mathrm{F}$ est le moment de Fermi, qui fixe le niveau de la surface de la mer, et la somme est restreinte aux excitations supérieures à k_F puisque les électrons de la paire ne peuvent être excités qu'à condition d'en sortir : c'est ici qu'intervient le Principe de Pauli, et de façon déterminante.

Il s'agit maintenant de résoudre l'équation aux valeurs propres, le Hamiltonien de la paire ayant la forme $H = H_1(\vec{r}_1) + H_1(\vec{r}_2) + V(\vec{r}_1, \vec{r}_2)$, où H_1 est le Hamiltonien pour un électron, et V l'interaction entre les deux électrons. En reportant l'expression (32.66)

[25] *A contrario*, à une dimension d'espace, un puits carré a toujours au moins un état lié, aussi peu profond soit-il (Tome I, sous-section 15.3.1).

dans l'équation $H\psi_0 = E\psi_0$, et en identifiant les composantes de Fourier, on obtient :

$$2\varepsilon^{(0)}(\vec{k})\Phi_{\vec{0}}(\vec{k}) + \sum_{\vec{k}',\, \|\vec{k}'\|>k_\mathrm{F}} \mathcal{V}(\vec{k}-\vec{k}')\Phi(\vec{k}') = E\Phi_{\vec{0}}(\vec{k}) \ ; \qquad (32.67)$$

dans cette équation $\mathcal{V}(\vec{k}-\vec{k}')$ est la transformée de Fourier $L^{-3/2}\int_{\mathbb{R}^3} V(\vec{r})\mathrm{e}^{-\mathrm{i}(\vec{k}-\vec{k}').\vec{r}}\,\mathrm{d}^3r$, L étant le côté de la (grande) boîte cubique de confinement. Comme auparavant, $\varepsilon^{(0)}(\vec{k})$ est la relation de dispersion d'un électron libre.

La discussion générale de l'équation (32.67) n'est guère possible ; afin de mettre en évidence très simplement l'existence d'une paire liée, Cooper choisit de prendre :

$$\mathcal{V}(\vec{k}-\vec{k}') = \begin{cases} -V_0 & \forall\vec{k}',\ k_\mathrm{F} < \|\vec{k}'\| < k_\mathrm{max} \quad (V_0 > 0) \\ 0 & \forall\vec{k}',\ \|\vec{k}'\| > k_\mathrm{max} \end{cases} \qquad (32.68)$$

où k_max est un vecteur d'onde de coupure, associé à l'énergie $E_\mathrm{max} \stackrel{\text{déf}}{=} \varepsilon^{(0)}(k_\mathrm{max}) \gg E_\mathrm{F}$. Avec cette modélisation, (32.67) donne :

$$\Phi_{\vec{0}}(\vec{k}) = \frac{V_0}{2\varepsilon^{(0)}(\vec{k}) - E} S_{k_\mathrm{max}} \ , \qquad S_{k_\mathrm{max}} \stackrel{\text{déf}}{=} \sum_{\vec{k}',\, k_\mathrm{F}<\|\vec{k}'\|<k_\mathrm{max}} \Phi_{\vec{0}}(\vec{k}') \ ; \qquad (32.69)$$

reportant alors l'expression de gauche de $\Phi_{\vec{0}}(\vec{k})$ dans la définition de la somme S_{k_max} à droite, et simplifiant par le facteur homogène S_{k_max}, on obtient :

$$\boxed{\sum_{\vec{k}',\, k_\mathrm{F}<\|\vec{k}'\|<k_\mathrm{max}} \frac{1}{2\varepsilon^{(0)}(\vec{k}') - E} = \frac{1}{V_0}} \qquad (32.70)$$

À nouveau, et comme pour les magnons (voir éq. (30.231)), il s'agit ni plus ni moins que d'une équation fixant les valeurs propres E. Avec les conditions aux limites habituelles, $\varepsilon^{(0)}(\vec{k}')$ est en réalité de la forme $\frac{\hbar^2}{2m}\vec{k}_n^2 \equiv \varepsilon_n^{(0)}$, où $\vec{k}_n$ est l'une des valeurs très denses imposées par la granulation dans l'espace réciproque induite par les conditions cycliques ; c'est pourquoi une écriture plus explicite de (32.70) est :

$$\boxed{\sum_{\vec{k}_n,\, k_\mathrm{F}<\|\vec{k}_n\|<k_\mathrm{max}} \frac{1}{2\varepsilon_n^{(0)} - E} = \frac{1}{V_0}} \qquad (32.71)$$

Les solutions de cette équation se représentent graphiquement comme indiqué sur la figure 32.4, en intersectant le graphe du premier membre de (32.71) avec la droite horizontale d'ordonnée $1/V_0$; on voit immédiatement que, *aussi petit* que soit V_0, il existe une solution d'énergie *inférieure* à $2E_\mathrm{F}$: c'est la paire liée de Cooper.

Tout comme pour les magnons, à partir du moment où on s'intéresse exclusivement à l'état lié situé sous le *continuum* démarrant à $2E_\mathrm{F}$, il n'y a aucun risque, dans (32.71), à

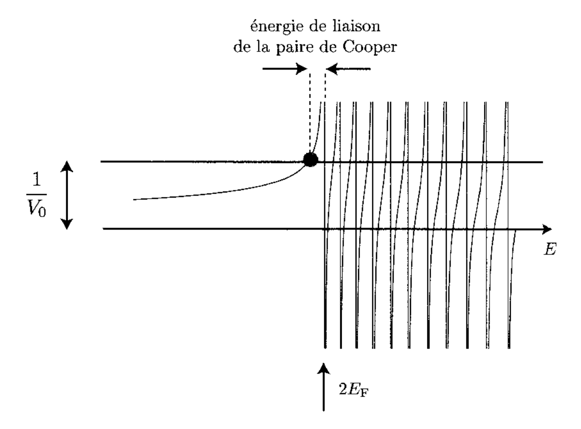


Figure 32.4: Représentation graphique schématique de l'équation aux valeurs propres (32.71) ; le cercle représente la solution d'énergie inférieure à $2E_F$ (paire de Cooper). La figure n'est pas à l'échelle : les arcs divergents sur la droite sont en fait infiniment serrés puisque l'énergie $\varepsilon^{(0)}(\vec{k}_n) \equiv \varepsilon_n^{(0)}$ varie de façon quasi-continue ; par ailleurs, les points d'intersection ont une ordonnée très élevée puisque $V_0 \ll E_F$.

remplacer la somme sur $\vec{k}_n$ par l'intégrale correspondante ; désignant par $\mathcal{D}(E)$ la densité d'états, on obtient :

$$\int_{E_F}^{E_F+E_{max}} \frac{\mathcal{D}(E')}{2E'-E}\,\mathrm{d}E' = \frac{1}{V_0} \ ; \tag{32.72}$$

comme $V_0 \ll E_F$, seul compte le *voisinage* du niveau de Fermi, ce qui autorise à substituer $\mathcal{D}(E')$, fonction lentement variable, par sa valeur en E_F. L'intégration est alors élémentaire et donne l'équation suivante pour l'énergie E_b de la paire liée :

$$\ln\frac{2E_F - E_b + 2E_{max}}{2E_F - E_b} = \frac{2}{V_0\mathcal{D}(E_F)} \ , \tag{32.73}$$

dont la solution est $E_b = 2E_F - 2E_{max}\left[e^{2/[V_0\mathcal{D}(E_F)]} - 1\right]^{-1}$. Comme usuellement $V_0\mathcal{D}(E_F)$ est assez petit ($\sim 0,1$), on retient généralement :

$$\boxed{E_b \simeq 2E_F - 2E_{max}\,e^{-2/[V_0\mathcal{D}(E_F)]}} \tag{32.74}$$

d'où l'énergie de liaison d'une paire :

$$E_{liaison} = 2E_{max}\,e^{-2/[V_0\mathcal{D}(E_F)]} \ . \tag{32.75}$$

Un tel résultat ne saurait être obtenu par une méthode perturbative : l'énergie de liaison est une fonction présentant une singularité essentielle en $V_0 = 0$, qui ne peut être le résultat d'un développement en puissances. Une fois trouvée la valeur propre E_b, la fonction d'onde peut être explicitement écrite ; avec (32.66) et $\Phi_{\vec{0}}$ donné en (32.69), on a :

$$\psi_0(\vec{r}_1, \vec{r}_2) = C \sum \frac{1}{2\varepsilon^{(0)}(\vec{k}) - E} \cos\left(\vec{k}.(\vec{r}_1 - \vec{r}_2)\right) \tag{32.76}$$

Comme E_b est juste au-dessous de $2E_F$, le poids est maximum pour les valeurs de $\vec{k}$ de module juste au-dessus de k_F, ce qui rassure sur le fait que la dépendance précise de $\mathcal{V}(\vec{k}-\vec{k}')$ aux grandes énergies ne joue pas un grand rôle. Par ailleurs, la fonction ψ_0 est visiblement à symétrie sphérique, puisqu'elle apparaît comme la transformée de Fourier d'une fonction ne dépendant que de $\|\vec{k}\,\|^2$: $\psi_0(\vec{r}_1, \vec{r}_2) \equiv \psi_0(\|\vec{r}_1 - \vec{r}_2\|)$; après intégration sur les angles, on obtient :

$$\psi_0(r) = \frac{C'}{r} \int_{k_F}^{k_{max}} \frac{k}{\frac{\hbar^2 k^2}{m} - E_b} \sin kr \, \mathrm{d}k \quad , \tag{32.77}$$

où l'intégrale est bien définie puisque $E_b < 2E_F \equiv \hbar^2 k_F^2/m$. Le facteur pondérant le sinus a une largeur $\Delta k \sim k_F - \hbar^{-1}\sqrt{E_b/m} \ll k_F$, de sorte que la largeur spatiale de la fonction d'onde de la paire est $\Delta r \gg 1/k_F$; comme k_F^{-1} est l'ordre de grandeur de la distance moyenne entre deux électrons de la mer de Fermi (voir (30.78)), on en déduit que si de nombreuses paires peuvent finalement se former, elles seront extraordinairement enchevêtrées les unes dans les autres.

Le résultat fondamental de Cooper est donc que la mer de Fermi *normale* est instable vis-à-vis de la formation de paires dès qu'il existe une interaction attractive entre les électrons, aussi petite soit-elle. Dès lors, une question se pose : une telle interaction est-elle possible ? Frölich [268] a été le premier à soupçonner que l'interaction entre les électrons et les phonons du métal joue un rôle de premier plan, une hypothèse largement confirmée par l'effet isotopique. Physiquement, l'idée est la suivante : un premier électron attire un ion du réseau qui, déplacé, laisse une lacune positive derrière lui ; une fois constituée, cette lacune est à son tour capable d'attirer un deuxième électron. Au total, ceci engendre bien une interaction effective *attractive* entre les deux électrons. La question de savoir si elle est suffisamment intense pour compenser la répulsion coulombienne est une question subtile, qui fait notamment intervenir la constante diélectrique du gaz d'électrons sans laquelle, évidemment, tout espoir serait vain d'obtenir une attraction nette attractive. De surcroît, la nécessité de comparaisons quantitatives serrées exige le calcul aussi précis que possible de la structure de bande et du couplage électron-phonon.

La théorie BCS, techniquement fort complexe, repose donc sur la possibilté de formation de paires de Cooper, admettant *par hypothèse* l'existence de l'interaction attractive. Les électrons d'une paire ont des spins opposés, tout comme le sont leurs impulsions ; formellement, l'état d'une telle paire isolée s'écrit :

$$|\phi_{\vec{k}}\rangle = c^\dagger_{\vec{k}\uparrow} c^\dagger_{-\vec{k}\downarrow} |\mathrm{vac}\rangle \quad , \tag{32.78}$$

où les $c^\dagger_{\vec{k}\uparrow/\downarrow}$ sont les opérateurs de création de fermions (qui anticommutent, voir (22.135)) et $|\mathrm{vac}\rangle$ est l'état vide (aucun électron) – ce n'est pas la mer de Fermi *normale*. L'état le plus général contenant une seule paire au plus est donc de la forme :

$$|\psi_{\vec{k}}\rangle = \big(u_{\vec{k}} + v_{\vec{k}}\, c^\dagger_{\vec{k}\uparrow} c^\dagger_{-\vec{k}\downarrow}\big) |\mathrm{vac}\rangle \quad , \tag{32.79}$$

où les coefficients $u_{\vec{k}}$ et $v_{\vec{k}}$, par leur modules carrés, donnent la probabilité d'observer l'état vide ou l'état à une paire. L'état $|\psi_{\vec{k}}\rangle$ est en réalité une *superposition* de ces deux

possibilités, qui porte en germe toute la cohérence quantique et ses potentialités. Pour décrire un état du métal contenant potentiellement un nombre arbitraire de paires, le trio BCS écrit la fonction d'onde fondamentale sous la forme :

$$|\Psi_{\rm g}\rangle = \prod_{\vec{k}} \big(u_{\vec{k}} + v_{\vec{k}}\, c^\dagger_{\vec{k}\uparrow} c^\dagger_{-\vec{k}\downarrow}\big)|\text{vac}\rangle \ , \tag{32.80}$$

où le produit sur $\vec{k}$ épuise un certain ensemble de valeurs de cardinal $\mathcal{N}$ si on borne le nombre de paires possibles. Trivialement, si tous les $v_{\vec{k}}$ sont nuls, c'est l'état vide ; à l'extrême opposé, si tous les $u_{\vec{k}}$ le sont, c'est la mer de Fermi ordinaire. L'état écrit en (32.80) semble donc posséder la flexibilité suffisante pour permettre le développement de l'instabilité de la mer de Fermi et la mise en évidence d'un état fondamental d'un nouveau type. Le traitement original de BCS consiste alors à utiliser la méthode variationnelle, qui requiert d'une part le choix d'une fonction d'essai contenant des paramètres – elle est de la forme (32.80), les paramètres étant les $u_{\vec{k}}$ et les $v_{\vec{k}}$ –, d'autre part d'un Hamiltonien H.

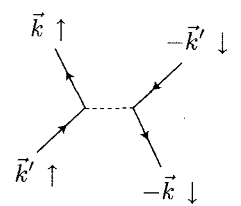


Figure 32.5: Représentation schématique des termes d'interaction retenus dans le Hamiltonien d'appariement (32.81).

Afin de simplifier un problème d'une immense complexité, le choix fait pour H est le suivant[26] :

$$H = \sum_{\vec{k},s=\uparrow,\downarrow} \varepsilon^{(0)}(\vec{k}) + \sum_{\vec{k},\vec{k}'} \mathcal{V}(\vec{k}-\vec{k}')c^\dagger_{\vec{k}\uparrow} c^\dagger_{-\vec{k}\downarrow} c_{-\vec{k}'\downarrow} c_{\vec{k}'\uparrow} \ , \tag{32.81}$$

opérateur censé retenir les interactions principales responsables de l'apparition de la phase supraconductrice : le terme d'interaction entre électrons est restreint aux termes créant exclusivement des paires d'impulsion totale nulle et de spins opposés (voir fig. 32.5). Avec la modélisation de Cooper, (32.68), ceci permet d'abord de mettre en évidence un *gap* d'énergie fini noté $\Delta(0)$:

$$\Delta(0) = 2E_{\rm max}\, {\rm e}^{-1/[\mathcal{D}(E_{\rm F})V_0]} \ , \tag{32.82}$$

puis de trouver les coefficients $u_{\vec{k}}$ et $v_{\vec{k}}$, et enfin de calculer l'énergie de l'état fondamental ; le résultat majeur est qu'elle est de fait *inférieure* à celle de la mer de Fermi :

$$\boxed{E_{\rm supra} - E_{\rm mer\,de\,Fermi} = -2\mathcal{D}(E_{\rm F})E^2_{\rm max}\, {\rm e}^{-2/[\mathcal{D}(E_{\rm F})V_0]}} \tag{32.83}$$

[26]Voir (22.170).

Ce résultat se comprend bien physiquement ; la différence d'énergie s'écrit (voir (32.75)) $\frac{1}{2}[\mathcal{D}(E_{\rm F})E_{\rm liaison}]\,E_{\rm liaison}$. $\mathcal{D}(E_{\rm F})E_{\rm liaison}$ est le nombre d'électrons essentiellement concernés par la possibilité d'appariement (le facteur $\frac{1}{2}$ évite de compter deux fois une même paire), et chaque paire constituée abaisse l'énergie de $E_{\rm liaison}$. L'apparition du *gap* d'énergie $\Delta(0)$ est prometteur puisqu'il est avéré expérimentalement que la chaleur spécifique a, pour $T < T_{\rm c}$, un comportement exponentiel du genre $\propto {\rm e}^{-aT_{\rm c}/T}$; à ce stade, $\Delta(0)$ semble devoir être considéré comme la valeur à $T = 0$ d'un *gap* dépendant de la température, $\Delta(T)$, dont l'annulation pour une certaine température $T_{\rm c}$ signerait la transition supra $\rightarrow$ normal. Une fois obtenus ces résultats, l'étape suivante est tout naturellement l'étude à température finie.

Le cadre variationnel choisi originalement dans la théorie BCS fournit un scénario permettant de démontrer l'existence d'un état fondamental supraconducteur, mais ne donne, par construction, aucune information précise sur le comportement du métal à température *finie*, lequel exige de connaître les premières excitations. Compte tenu de l'aspect problème à N-corps, avec $N \gg 1$, il est légitime de penser que des méthodes de type champ moyen astucieuses vont permettre de progresser, l'idée étant que pour un problème où tant de particules sont concernées simultanément, les *fluctuations* sont (relativement) faibles. Techniquement, une manière de procéder consiste à manipuler les opérateurs centrés $c_{-\vec{k}\downarrow}c_{\vec{k}\uparrow} - \gamma_{\vec{k}}$ où $\gamma_{\vec{k}} \stackrel{\rm déf}{=} \langle c_{-\vec{k}\downarrow}c_{\vec{k}\uparrow}\rangle$, et à *linéariser* le terme d'interaction en délaissant les quantités quadratiques par rapport à l'écart à la valeur moyenne ; très précisément, on commence par écrire :

$$c_{-\vec{k}\downarrow}c_{\vec{k}\uparrow} = \gamma_{\vec{k}} + (c_{-\vec{k}\downarrow}c_{\vec{k}\uparrow} - \gamma_{\vec{k}}) \equiv \gamma_{\vec{k}} + \delta_{\vec{k}} \ , \tag{32.84}$$

puis on reporte dans (32.81) en laissant tomber les termes quadratiques en $\delta_{\vec{k}}$. Ce faisant, on obtient un Hamiltonien linéarisé :

$$H_{\rm lin} = \sum_{\vec{k},s=\uparrow,\downarrow} \varepsilon_{\vec{k}} + \sum_{\vec{k},\vec{k}'} \mathcal{V}(\vec{k}-\vec{k}')(\gamma_{\vec{k}'}c^\dagger_{\vec{k}\uparrow}c^\dagger_{-\vec{k}\downarrow} + \gamma^*_{\vec{k}}c_{-\vec{k}'\downarrow}c_{\vec{k}'\uparrow} - \gamma^*_{\vec{k}}\gamma_{\vec{k}'}) \ , \tag{32.85}$$

où $\varepsilon_{\vec{k}} \stackrel{\rm déf}{=} \varepsilon^{(0)}(\vec{k}) - E_{\rm F}$, et où les moyennes $\gamma_{\vec{k}}$ devront être en temps utile trouvées par une équation *fermant* la procédure de linéarisation[27] :

$$\boxed{\gamma_{\vec{k}} = \langle c_{-\vec{k}\downarrow}c_{\vec{k}\uparrow}\rangle} \tag{32.86}$$

en calculant le second membre avec la solution une fois trouvée ; celle-ci étant elle-même paramétrée par les $\gamma_{\vec{k}}$, la résolution de cette équation achèvera la solution.

Après cette linéarisation, on peut effectuer l'une des sommations dans le terme d'interaction de l'expression (32.85) ; dans des notations conventionnelles, on définit :

$$\Delta_{\vec{k}} \stackrel{\rm déf}{=} \sum_{\vec{k}'} \mathcal{V}(\vec{k}-\vec{k}')\gamma_{\vec{k}'} \ , \tag{32.87}$$

[27]L'équation (32.86) est strictement de même nature que celle qui, en théorie de Weiss du ferromagnétisme – théorie de champ moyen par excellence –, se traduit par $\mathcal{M} = \tanh(\beta J\mathcal{M} + \mathcal{B})$ et permet de montrer simplement l'existence d'un changement de phase à une température $T_{\rm c}$ telle que $\beta_{\rm c}J = 1$ (voir par exemple [24]).

donnant la forme suivante pour H_{lin} :

$$\boxed{H_{\text{lin}} = \sum_{\vec{k},s=\uparrow,\downarrow} \varepsilon_{\vec{k}} + \sum_{\vec{k}} (\Delta_{\vec{k}}\, c^\dagger_{\vec{k}\uparrow} c^\dagger_{-\vec{k}\downarrow} + \Delta^*_{\vec{k}}\, c_{-\vec{k}\downarrow} c_{\vec{k}\uparrow} - \Delta_{\vec{k}} \gamma^*_{\vec{k}})} \tag{32.88}$$

Comme l'ont montré Bogoliubov [269] et Valatin [270], il existe une procédure simple permettant de diagonaliser tout Hamiltonien de cette forme, le point essentiel étant qu'il est *quadratique* par rapport à des opérateurs de création/annihilation. L'idée est de former les bonnes combinaisons linéaires de ces opérateurs qui annulent les termes non-diagonaux, ce qui définit de nouveaux modes propres $A_{\vec{k}}$ et $B_{\vec{k}}$. Tirant parti d'emblée du fait que la transformation doit être unitaire, on pose ainsi *a priori* :

$$A_{\vec{k}} \stackrel{\text{déf}}{=} \cos\theta_{\vec{k}}\, c_{\vec{k}\uparrow} - \mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \sin\theta_{\vec{k}}\, c^\dagger_{-\vec{k}\downarrow} \ , \qquad B^\dagger_{\vec{k}} \stackrel{\text{déf}}{=} \mathrm{e}^{-\mathrm{i}\phi_{\vec{k}}} \sin\theta_{\vec{k}}\, c_{\vec{k}\uparrow} + \cos\theta_{\vec{k}}\, c^\dagger_{-\vec{k}\downarrow} \ , \tag{32.89}$$

les angles $\theta_{\vec{k}}$ et $\phi_{\vec{k}}$ étant à déterminer. $A_{\vec{k}}$ détruit une particule $(\vec{k}\uparrow)$ *ou* crée une particule $(-\vec{k}\downarrow)$; dans tous les cas, l'impulsion du système varie de $-\vec{k}$, et le spin diminue de $\frac{\hbar}{2}$; les variations étant dans le sens opposé en ce qui concerne $B_{\vec{k}}$. Comme la transformation est unitaire, les relations de commutation sont préservées :

$$\{A_{\vec{k}},\, A^\dagger_{\vec{k}'}\} = \{B_{\vec{k}},\, B^\dagger_{\vec{k}'}\} = \delta_{\vec{k}\vec{k}'}\mathbf{1} \ , \qquad \{A_{\vec{k}},\, B^\dagger_{\vec{k}'}\} = 0 \ , \qquad \text{etc.} \tag{32.90}$$

La transformation inverse de (32.89) est :

$$c_{\vec{k}\uparrow} = \cos\theta_{\vec{k}}\, A_{\vec{k}} + \mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \sin\theta_{\vec{k}}\, B^\dagger_{\vec{k}} \ , \qquad c^\dagger_{-\vec{k}\downarrow} = -\mathrm{e}^{-\mathrm{i}\phi_{\vec{k}}} \sin\theta_{\vec{k}}\, A_{\vec{k}} + \cos\theta_{\vec{k}}\, B^\dagger_{\vec{k}} \ . \tag{32.91}$$

Le report des expressions (32.91) dans (32.88) produit quatre types de termes : $A^\dagger_{\vec{k}} A_{\vec{k}}$, $B^\dagger_{\vec{k}} B_{\vec{k}}$, $A^\dagger_{\vec{k}} B^\dagger_{\vec{k}}$ et $A_{\vec{k}} B_{\vec{k}}$, et on constate qu'il est possible d'annuler les coefficients de ces deux derniers produits, à condition de choisir les angles $\theta_{\vec{k}}$ et $\phi_{\vec{k}}$ satisfaisant :

$$(\Delta^*_{\vec{k}}\, \mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \tan\theta_{\vec{k}})^2 - 2\varepsilon_{\vec{k}} \Delta^*_{\vec{k}} \mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \tan\theta_{\vec{k}} - |\Delta_{\vec{k}}|^2 = 0 \ ; \tag{32.92}$$

cette équation pour $X \stackrel{\text{déf}}{=} \Delta^*_{\vec{k}}\, \mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \tan\theta_{\vec{k}}$ donne $X = \varepsilon_{\vec{k}} - \left(\varepsilon^2_{\vec{k}} + |\Delta_{\vec{k}}|^2\right)^{1/2}$, le choix du signe assurant d'avoir la solution la plus basse en énergie ; on en déduit :

$$\cos 2\theta_{\vec{k}} = \frac{\varepsilon_{\vec{k}}}{2\sqrt{\varepsilon^2_{\vec{k}} + |\Delta_{\vec{k}}|^2}} \ . \tag{32.93}$$

Avec (32.92), on a $\mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \sin\theta_{\vec{k}} \cos\theta_{\vec{k}} = \frac{1}{2\varepsilon_{\vec{k}}}(\Delta^*_{\vec{k}} \mathrm{e}^{2\mathrm{i}\phi} \sin^2\theta_{\vec{k}} - \Delta_{\vec{k}} \cos^2\theta_{\vec{k}})$; comme X est réel, la phase $\phi_{\vec{k}}$ est égale à l'argument de $\Delta_{\vec{k}}$, de sorte que :

$$\mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \sin\theta_{\vec{k}} \cos\theta_{\vec{k}} = -\frac{\Delta_{\vec{k}}}{2\sqrt{\varepsilon^2_{\vec{k}} + |\Delta_{\vec{k}}|^2}} \ . \tag{32.94}$$

Les coefficients de la transformation de Bogoliubov - Valatin étant trouvés, le Hamiltonien linéarisé prend par construction la forme canonique :

$$H_{\text{lin}} = \sum_{\vec{k}} \left[\varepsilon_{\vec{k}} - E(\vec{k}) - \Delta_{\vec{k}} \gamma^*_{\vec{k}}\right] + \sum_{\vec{k}} E(\vec{k})(A^\dagger_{\vec{k}} A_{\vec{k}} + B^\dagger_{\vec{k}} B_{\vec{k}}) \tag{32.95}$$

Le premier terme est une constante, qui se trouve être strictement égale à l'énergie de l'état fondamental BCS ; le second terme décrit les nouveaux modes normaux après diagonalisation. Considérant l'état fondamental BCS comme le nouveau *vide* du métal à température nulle, les $A_{\vec{k}}$ et $B_{\vec{k}}$ représentent les *quasi-particules* du système, premières excitations au-dessus de l'état vide. Elles ont l'énergie $E(\vec{k})$:

$$\boxed{E(\vec{k}) = \sqrt{\varepsilon_{\vec{k}}^2 + |\Delta_{\vec{k}}|^2}} \tag{32.96}$$

puisque $E(\vec{k}) \geq |\Delta_{\vec{k}}| > 0$, ces excitations élémentaires présentent un *gap* d'énergie relativement à l'état vide ; elles sont souvent appelées *Bogoliubons.*

Il reste maintenant à expliciter l'équation d'auto-cohérence (32.86), qui va montrer que le *gap* est de fait fonction de la température T. Revenant à la transformation (32.91), la valeur moyenne cherchée s'exprime comme :

$$\langle c_{-\vec{k}\downarrow} c_{\vec{k}\uparrow} \rangle = \langle (-\mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \sin\theta_{\vec{k}} A_{\vec{k}}^\dagger + \cos\theta_{\vec{k}} B_{\vec{k}})(\cos\theta_{\vec{k}} A_{\vec{k}} + \mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \sin\theta_{\vec{k}} B_{\vec{k}}^\dagger)\rangle \ ; \tag{32.97}$$

les termes $A_{\vec{k}}^\dagger B_{\vec{k}}^\dagger$ et $A_{\vec{k}} B_{\vec{k}}$ ont tous leurs éléments diagonaux nuls et ne contribuent donc pas à la moyenne ; en utilisant $B_{\vec{k}} B_{\vec{k}}^\dagger = \mathbf{1} - B_{\vec{k}}^\dagger B_{\vec{k}}$, on obtient :

$$\gamma_{\vec{k}} \equiv \langle c_{-\vec{k}\downarrow} c_{\vec{k}\uparrow} \rangle = \sin\theta_{\vec{k}} \cos\theta_{\vec{k}}\, \mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \big(1 - \langle A_{\vec{k}}^\dagger A_{\vec{k}} + B_{\vec{k}}^\dagger B_{\vec{k}} \rangle\big) \ ; \tag{32.98}$$

si $f(E)$ désigne la distribution de Fermi $(\mathrm{e}^{\beta E} + 1)^{-1}$, $\langle A_{\vec{k}}^\dagger A_{\vec{k}} \rangle = \langle B_{\vec{k}}^\dagger B_{\vec{k}} \rangle = f\big(E(\vec{k})\big)$, de sorte que $1 - \langle A_{\vec{k}}^\dagger A_{\vec{k}} + B_{\vec{k}}^\dagger B_{\vec{k}} \rangle = 1 - 2f\big(E(\vec{k})\big) = \tanh \frac{1}{2}\beta E(\vec{k})$, d'où l'expression de $\gamma_{\vec{k}}$ en fonction de $\Delta_{\vec{k}}$:

$$\gamma_{\vec{k}} = \sin\theta_{\vec{k}} \cos\theta_{\vec{k}}\, \mathrm{e}^{\mathrm{i}\phi_{\vec{k}}} \tanh\left(\frac{1}{2}\beta\sqrt{\varepsilon_{\vec{k}}^2 + \Delta_{\vec{k}}^2}\right) \ ; \tag{32.99}$$

revenant à la définition (32.87), et compte tenu de (32.94), on en déduit l'équation à résoudre pour clore le problème :

$$\Delta_{\vec{k}} = -\sum_{\vec{k}'} \mathcal{V}(\vec{k} - \vec{k}') \frac{\Delta_{\vec{k}'}}{2E(\vec{k}')} \tanh \frac{1}{2}\beta E(\vec{k}') \ . \tag{32.100}$$

La dépendance de $\Delta_{\vec{k}}$ bloque tout progrès ; adoptant maintenant la modélisation de Cooper (32.68), $\Delta_{\vec{k}}$ devient une constante Δ, et l'équation se simplifie en[28] :

$$\frac{1}{V_0} = \frac{1}{2}\sum_{\vec{k}} \frac{1}{E(\vec{k})} \tanh \frac{1}{2}\beta E(\vec{k}) \ , \tag{32.101}$$

l'énergie $E(\vec{k})$ étant donnée en (32.96) ; substituant l'intégrale à la somme, on obtient[29] :

$$\boxed{\frac{1}{V_0 \mathcal{D}(E_\mathrm{F})} = \int_0^{E_\mathrm{max}} \frac{1}{\sqrt{\varepsilon^2 + \Delta^2}} \tanh\left(\frac{1}{2}\beta\sqrt{\varepsilon^2 + \Delta^2}\right) \mathrm{d}\varepsilon} \tag{32.102}$$

[28] On laisse de côté la solution $\Delta = 0$, correspondant trivialement à l'état normal, dont on sait déjà qu'il est instable et a une énergie supérieure.

[29] On a utilisé la symétrie de $\varepsilon_{\vec{k}}$ de part et d'autre de E_F.

Ceci est l'étape finale de la procédure de champ moyen, et fournit une équation pour le *gap* ; sa résolution donne la variation de $\Delta(T)$ en température, un résultat majeur. Avant de continuer, notons le caractère *effectif* des modes propres fermioniques $A_{\vec{k}}$ et $B_{\vec{k}}$: étant des fonctions du *gap* $\Delta(T)$, ils sont paramétrés par la température. Une situation de nature comparable a été rencontrée à propos de la première correction quantique à la fonction de partition classique d'un gaz parfait (voir problème 22.6.6), où on a mis en évidence une interaction *effective* répulsive entre fermions (et attractive pour les bosons). Dans ces deux situations, c'est le Principe de Pauli qui est à l'œuvre ; pour la supraconductivité, cette dépendance paramétrique en température est une autre manifestation du rôle crucial joué par le bain d'électrons "normaux".

Discutons maintenant la dépendance du *gap* en température. À température nulle ($\beta = +\infty$), l'intégration de (32.102) est élémentaire et donne :

$$\Delta(T=0) = \frac{E_{\text{max}}}{\sinh[1/V_0\mathcal{D}(E_\text{F})]} \simeq 2E_{\text{max}}\,\mathrm{e}^{-1/[V_0\mathcal{D}(E_\text{F})]} \ , \tag{32.103}$$

soit essentiellement le résultat BCS (32.82), obtenu d'une façon finalement assez différente. Par ailleurs, l'égalité (32.102) montre que la solution $\Delta(T)$ *décroît* quand la température augmente : la température critique T_c est celle où le *gap* s'annule, le métal transitant à ce point précis dans l'état normal ; T_c est ainsi fixée par :

$$\frac{1}{V_0\mathcal{D}(E_\text{F})} = \int_0^{E_{\text{max}}} \frac{1}{\varepsilon}\tanh\frac{1}{2}\beta_\text{c}\varepsilon\,\mathrm{d}\varepsilon \iff \frac{1}{V_0\mathcal{D}(E_\text{F})} = \int_0^{x_{\text{max}}} \frac{1}{x}\tanh x\,\mathrm{d}x \ , \tag{32.104}$$

où $x_{\text{max}} = E_{\text{max}}/(2k_\text{B}T_\text{c})$; l'intégrale de droite vaut :

$$\ln x_{\text{max}}\,\tanh x_{\text{max}} - \int_0^{x_{\text{max}}} \frac{\ln x}{\cosh^2 x}\mathrm{d}x \ . \tag{32.105}$$

La situation dite *de couplage faible* correspond au cas où $E_{\text{max}} \gg k_\text{B}T_\text{c}$; dans ces conditions, le terme tout intégré est exponentiellement voisin de $\ln x_{\text{max}}$; quant à l'intégrale résiduelle, elle est aussi exponentiellement proche de la même intégrale prise de 0 à l'infini, laquelle peut s'obtenir comme la limite[30] $\mu \to 0$ de $\frac{\mathrm{d}}{\mathrm{d}\mu}I(\mu)$, avec :

$$I(\mu) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{x^\mu}{\cosh^2 x}\,\mathrm{d}x \ . \tag{32.107}$$

Cette dernière intégrale se calcule en tournant dans $\mathbb{C}$, longeant de part et d'autre la coupure de la fonction multiforme x^μ, ce qui fait ramasser tous les résidus aux pôles

[30] Cette astuce repose sur l'égalité :

$$\ln Z = \lim_{\alpha\to 0} \frac{Z^\alpha - 1}{\alpha} \ ; \tag{32.106}$$

les notations ne sont pas innocentes : ce *truc* fut le point de départ d'une méthode d'approximation devenue célèbre, due à Mézard et Parisi [271], et appelée *méthode des répliques* (*replica trick*). Elle est fréquemment utilisée pour traiter les systèmes désordonnés, notamment les verres de spins, afin d'en savoir plus sur leur fonction de partition Z.

(d'ordre deux) de $(\cosh z)^{-2}$ en $z_n = (2n+1)\mathrm{i}\frac{\pi}{2}$; on trouve finalement[31] :

$$I(\mu) = \mu\pi^{\mu}(1-2^{1-\mu})\frac{\zeta(1-\mu)}{\cos\frac{\pi\mu}{2}} = 2^{1-\mu}(1-2^{1-\mu})\Gamma(\mu+1)\zeta(\mu) \ , \qquad (32.108)$$

où $\zeta(z)$ est la fonction de Riemann. Connaissant les valeurs :

$$\zeta(0) = -\frac{1}{2} \ , \qquad \zeta'(0) = -\frac{1}{2}\ln 2\pi \ , \qquad \Gamma'(1) = -\mathrm{C} \simeq -0,577 \ , \qquad (32.109)$$

on trouve finalement $\int_0^{+\infty} \frac{\ln x}{\cosh^2 x}\mathrm{d}x = \ln\frac{\pi}{4} - \mathrm{C} \simeq -0,819$. Ceci fait, l'égalité (32.104) s'écrit $V_0\mathcal{D}(E_\mathrm{F})\ln\left(\frac{4\mathrm{e}^\mathrm{C}}{\pi}x_\mathrm{max}\right) = 1$, d'où la température critique :

$$\boxed{k_\mathrm{B}T_\mathrm{c} = \frac{2\mathrm{e}^\mathrm{C}}{\pi}\,E_\mathrm{max}\,\mathrm{e}^{-1/[V_0\mathcal{D}(E_\mathrm{F})]} \simeq 1,1\,E_\mathrm{max}\,\mathrm{e}^{-1/[V_0\mathcal{D}(E_\mathrm{F})]}} \qquad (32.110)$$

La comparaison avec (32.82) montre que $\Delta(0) \simeq 1,8\,k_\mathrm{B}T_\mathrm{c}$, qui se trouve en bon accord avec un grand nombre de données expérimentales (le plus souvent, on trouve un nombre compris entre $1,5$ et $2,5$).

Dans l'expression de la température critique (32.110), l'énergie de coupure E_max est en facteur, et pilote donc pour une grande part la valeur de T_c. Admettant le mécanisme d'appariement des électrons grâce à la médiation des phonons, cette énergie est de l'ordre de 10^{11} Hz, donnant $T_\mathrm{c} \sim 5$ à 10 K, d'où le *dogme* historique de la supraconductivité conventionnelle : ce sont les propres agents responsables de la supraconductivité qui en limitent la température critique à des valeurs extrêmement basses. L'intervention d'autres excitations élémentaires dans le rôle des phonons (excitons[32] par exemple, [272]) a donné certains espoirs qui n'ont pas été confirmés à ce jour.

[31] La deuxième égalité provient de l'une des relations fonctionnelles pour $\zeta(z)$ ([38], § 9.535.3), $\Gamma(z)$ étant toujours la fonction d'Euler de deuxième espèce. Par ailleurs, C est la constante d'Euler ([38], p. xxviii).

[32] Les *excitons* sont des excitations élémentaires subtiles mises en évidence par Frenkel d'une part, par Mott et Wannier d'autre part.

L'exciton dit *de Frenkel* correspond primitivement à un *défaut* cristallin d'une nature très particulière. Soit par exemple un cristal moléculaire : un "défaut" peut être conceptuellement créé en imaginant que l'une des molécules a été portée dans un état excité, toutes les autres restant au fondamental. Cette situation ne correspond sûrement pas à un état propre, en raison de l'interaction (faible) des molécules entre elles, mais il existe des états de Bloch formés en combinaison linéaire de toutes les excitations locales – visiblement, le schéma procède exactement comme pour l'approximation des liaisons fortes (voir chapitre 30, sous-section 30.4.2), ou comme pour les magnons (section 30.6). Par une excitation à spectre assez large, on peut de fait construire un paquet d'ondes de ces états de Bloch pour préparer un état non-stationnaire où l'excitation est réellement localisée sur un petit nombre de molécules. Son déplacement ultérieur dans le cristal est rapide, comparé à celui d'un vrai défaut d'origine chimique ou structurelle, et est l'un des mécanismes de transfert d'énergie dans ces matériaux.

La formation de l'exciton dit *de Mott - Wannier* relève d'un scénario différent, où l'on se représente le fondamental d'un isolant (ou d'un semi-conducteur) comme formant le *vide* de ce système, les bandes de valence et de conduction étant respectivement pleine et vide. Dans un schéma à une particule, l'excitation ne peut se produire que moyennant une énergie au moins égale au *gap*, définissant le seuil de photoconductivité déjà évoqué (voir Remarque 2, p. 1411) ; cela fait, on se retrouve avec un électron dans la bande de conduction et un *trou* dans la bande de valence, celui-ci se comportant comme une lacune positive (voir par exemple le commentaire suivant (32.21)). Maintenant, si au lieu de considérer des états d'impulsion définie, on construit des paquets d'ondes donnant des états fortement localisés, il

Toujours dans la situation de couplage faible, $k_B T_c \ll E_{max}$, l'équation (32.102) donnant le *gap* $\Delta(T)$ peut, à des termes exponentiellement petits près, être mise sous la forme ([174], et problème 32.6.1) :

$$\ln \frac{\Delta(0)}{\Delta(T)} = 2 \int_0^{E_{max}} \frac{1}{\sqrt{\varepsilon^2 + [\Delta(T)]^2}} \frac{1}{e^{\beta\sqrt{\varepsilon^2+[\Delta(T)]^2}} + 1} d\varepsilon \ . \tag{32.111}$$

Les résultats précédents permettent de trouver la variation approximative du *gap* aux extrémités de l'intervalle $[0,\, T_c]$ (les calculs détaillés sont proposés dans le problème 32.6.1). Pour $T \lesssim T_c$, on trouve [174] :

$$\ln \frac{T}{T_c} + \frac{7}{8}\zeta(3)\Big(\frac{\Delta(T)}{\pi k_B T}\Big)^2 + \mathcal{O}\Big(\frac{\Delta}{k_B T}\Big)^4 = 0 \ , \tag{32.112}$$

d'où l'on tire :

$$T \lesssim T_c \ : \ \Delta(T) \simeq k_B T_c \Big(\frac{8\pi^2}{7\zeta(3)}\Big)^{1/2} \Big(1 - \frac{T}{T_c}\Big)^{1/2} \simeq 3{,}063\, k_B T_c \Big(1 - \frac{T}{T_c}\Big)^{1/2} \ , \tag{32.113}$$

l'exposant $\frac{1}{2}$ étant caractéristique des théories de champ moyen[33]. À l'autre extrémité, soit $T \ll T_c$, le calcul donne :

$$T \ll T_c \ : \ \Delta(T) \simeq \Delta(0)\Big[1 - \Big(\frac{2\pi k_B T}{\Delta(0)}\Big)^{1/2} e^{-\frac{\Delta(0)}{k_B T}}\Big] \ . \tag{32.114}$$

Au total, le *gap* varie comme indiqué sur la figure 32.6.

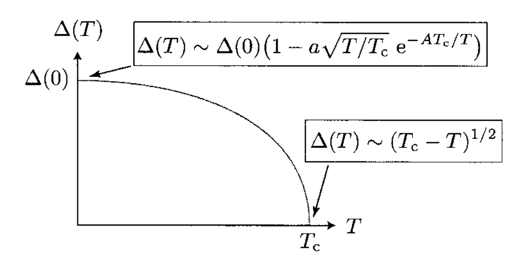


Figure 32.6: Fonction universelle donnant la variation du *gap* en fonction de la température pour la situation de couplage faible.

n'est plus possible d'omettre l'interaction (attractive !) entre l'électron et le trou, qui forment de fait un système hydrogénoïde (de charge totale nulle) susceptible d'exister dans des états liés : dans cette conception, l'exciton est une paire liée électron-trou, dont la dimension spatiale est très grande comparée à celle de l'exciton de Frenkel. La signature de ces excitons est l'apparition, dans le spectre d'excitation, de résonances fines situées évidemment *au-dessous* du seuil de photoconductivité, la différence étant précisément l'énergie de liaison de la paire.

[33]On trouve le même exposant pour l'annulation en T_c de la magnétisation spontanée d'un ferromagnétique traité en champ moyen (théorie de Weiss, modèle d'Ising), ou pour l'annulation de la différence de densité entre un liquide et un gaz en théorie de van der Waals [24]. De façon universelle, c'est l'exposant qui apparaît dans la théorie phénoménologique de Landau décrivant le voisinage du point critique.

Ainsi, la théorie BCS, fondamentalement microscopique, est capable d'expliquer le phénomène de base de la supraconductivité dans les situations où ses inévitables hypothèses simplificatrices sont légitimes. L'accord numérique est même extraordinairement bon, comme en attestent les valeurs théorique et expérimentales du rapport $\Delta(0)/(k_\mathrm{B}T_\mathrm{c})$, précisées juste après (32.110). Toutefois, elle ne constitue pas un cadre utilisable dans toutes les situations physiques rencontrées. En particulier, elle ne considère aucune inhomogénéité spatiale, dont la prise en compte dans le cadre BCS conduit à des problèmes quasi-inextricables, qu'il s'agisse de décrire des supraconducteurs dits *sales*, ou de l'interface entre deux métaux, l'un dans l'état supraconducteur, l'autre dans l'état normal. En pareilles circonstances, on lui préfère le plus souvent la théorie phénoménologique (et littéralement phénoménale) de Ginzburg - Landau, construite au tout début des années cinquante, dont les succès sont aussi prodigieux que ceux de la théorie de... Landau pour les phénomènes critiques.

32.6 Exercices et problèmes

32.6.1 Variation en température du *gap* supraconducteur (couplage faible)

1. Comportement à basse température :

   (a) Établir l'équation (32.111).

   (b) Simplifier cette équation dans la limite $T \ll T_\mathrm{c}$.

   (c) Par un changement de variable dans l'intégrale, montrer que l'équation du *gap* s'écrit :

   $$\ln\frac{\Delta(0)}{\Delta(T)} \simeq 2K_0\big(\beta\Delta(T)\big) \tag{32.115}$$

   où $K_0(x) \stackrel{\text{déf}}{=} \int_0^{+\infty} \mathrm{e}^{-x\cosh t}\,\mathrm{d}t$ est une certaine fonction de Bessel.

   (d) Sachant que pour $x \gg 1$, $K_0(x) \sim \sqrt{\frac{\pi}{2x}}\,\mathrm{e}^{-x}$, établir l'expression à très basse température du *gap* (voir (32.114)).

2. Pour trouver le comportement à $T \lesssim T_\mathrm{c}$, on revient à la forme (32.102).

   (a) On pose $\delta(T) \stackrel{\text{déf}}{=} \frac{1}{2}\beta\Delta(T)$. Montrer que l'équation s'écrit :

   $$\int_{\frac{1}{2}\beta_\mathrm{c}E_\mathrm{max}}^{\frac{1}{2}\beta E_\mathrm{max}} \frac{\tanh\sqrt{x^2+\delta^2}}{\sqrt{x^2+\delta^2}}\,\mathrm{d}x = \int_0^{\frac{1}{2}\beta_\mathrm{c}E_\mathrm{max}} \left(\frac{\tanh x}{x} - \frac{\tanh\sqrt{x^2+\delta^2}}{\sqrt{x^2+\delta^2}}\right)\mathrm{d}x \tag{32.116}$$

   (b) En déduire que l'on a :

   $$\ln\frac{T}{T_\mathrm{c}} \simeq \mathcal{I}\,\delta^2\ , \tag{32.117}$$

   où $\mathcal{I}$ est une certaine intégrale, dont on vérifiera qu'elle est finie.

(c) Montrer que ce résultat permet de retrouver le type de comportement (32.113) de $\Delta(T)$ pour T juste au-dessous de T_c.

(d) En effectuant plusieurs intégrations par parties, montrer que l'intégrale $\mathcal{I}$ est égale à $-\frac{1}{2}\int_0^{+\infty} \frac{\sinh x}{x \cosh^3 x}\,dx$, que l'on calcule sans peine par résidus. En déduire l'expression (32.113).

3. Autre petit exercice avec les fonctions analytiques : établir les égalités (32.108) en suivant la procédure indiquée.

32.6.2 Grandeurs thermodynamiques dans les phases normale et supraconductrice

Les résultats de la théorie BCS linéarisée font donc apparaître des modes fermioniques effectifs d'énergie $E(k)$ donnée en (32.96). La probabilité d'occupation de ces modes est la distribution de Fermi $f(E) = \frac{e^{-\beta E}}{1+e^{-\beta E}} = (e^{\beta E}+1)^{-1}$, et on note $f_{\vec{k}} \stackrel{\text{déf}}{=} f(E(\vec{k}))$.

Les résultats obtenus ci-dessous présentent des caractères d'universalité (notamment le saut de chaleur spécifique), que l'on retrouve dans la théorie de Landau des phénomènes critiques et, d'une façon générale, dans les traitements de champ moyen (par exemple : traitement de Weiss du ferromagnétisme [24]) ; tous donnent aussi un exposant $\frac{1}{2}$ pour le *paramètre d'ordre* (voir (32.113)).

1. Écrire l'expression en fonction des $f_{\vec{k}}$ de l'entropie statistique de la phase supraconductrice, S_{supra} ; la comparer à celle de la phase normale S_{norm} et les représenter graphiquement en fonction de la température.

2. En déduire que la chaleur spécifique de la phase *supra* est :

$$C_{\text{supra}} = -2\beta k_B \sum_{\vec{k}} \frac{\partial f_{\vec{k}}}{\partial E(\vec{k})}\left(E(\vec{k})^2 + \frac{1}{2}\beta\frac{d\Delta^2}{d\beta}\right) . \tag{32.118}$$

3. La chaleur spécifique dans la phase normale est $C_{\text{norm}} = \gamma T \equiv \frac{2\pi^2}{3}\mathcal{D}(E_F)k_B^2 T$, γ étant la constante de Sommerfeld ; on pose $\Delta C \stackrel{\text{déf}}{=} C_{\text{supra}} - C_{\text{norm}}$; montrer que ΔC a un saut en T_c, donné par :

$$\Delta C\big|_{T_c} = -\mathcal{D}(E_F)\frac{d\Delta^2}{dT} . \tag{32.119}$$

4. Une fois la chaleur spécifique obtenue, l'énergie interne U peut être obtenue par intégration. La chaleur spécifique étant finie en T_c, les deux énergies U_{supra} et U_{norm} sont égales en T_c. En déduire l'égalité :

$$U_{\text{supra}}(T) = U_{\text{norm}}(0) + \frac{1}{2}\gamma T_c^2 + \int_{T_c}^{T} C_{\text{supra}}(T)\,dT . \tag{32.120}$$

Illustrer ce résultat par un graphique.

5. Tracer les graphes en fonction de la température des énergies libres $F \stackrel{\text{déf}}{=} U - TS$ des deux phases.

Bibliographie

[1] Pierre CURIE, *"Sur la symétrie dans les phénomènes physiques, symétrie d'un champ électrique et d'un champ magnétique"*, Journal de Physique, **3**, 393 (1894)

[2] Claude COHEN-TANNOUDJI, Notes des cours du Collège de France, disponibles à `http://www.phys.ens.fr/cours/`

[3] Claude COHEN-TANNOUDJI, Jacques DUPONT-ROC et Gilbert GRYNBERG, *Photons et atomes - Introduction à l'électrodynamique quantique* (EDP Sciences, Paris, 2001)

[4] Eugen MERZBACHER, *Quantum Mechanics* (Wiley, New York, 1970)

[5] Marie-Antoinette TONNELAT, *Histoire du principe de relativité* (Flammarion, Paris, 1971)

[6] Erwin SCHRÖDINGER, *Physique quantique et représentation du monde* (collection Points-Sciences, Paris, 1992)

[7] `http://galileo.phys.virginia.edu/classes/252/michelson.html`

[8] Yvan SIMON, *Relativité restreinte* (Vuibert, Paris, 2004)

[9] Henri POINCARÉ, *La science et l'hypothèse* (Flammarion, Paris, 1968)

[10] Lev LANDAU et Evguéni LIFSHITZ, *Mécanique* (Mir, Moscou, 1966)

[11] Emmy NOETHER, *"Invariante Variationsprobleme"*, Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse, 235, (1918)

[12] Bertrand DELAMOTTE, *"Un soupçon de théorie des groupes : groupe des rotations et groupe de Poincaré"*, `www.lpthe.jussieu.fr/DEA/delamotte.html`

[13] Mikhaïl LAVRENTIEV et Boris CHABAT, *Méthodes de la théorie des fonctions d'une variable complexe* (Éditions Mir, Moscou, 1972)

[14] Eugen Paul WIGNER, *Gruppentheorie*, p. 251-254 (Frederick Wieweg und Sohn, Braunschweig, 1931)

[15] Albert MESSIAH, *Mécanique Quantique*, tome 2 (Dunod, Paris 1965)

[16] Eugen Paul WIGNER, *"The Unreasonable Effectiveness of Mathematics in the Natural Sciences"*, Comm. Pure and Appl. Math., **13**, n° 1 (1960)

[17] Vladimir Ivanovich SMIRNOV, *A course of Higher Mathematics. I* (Pergamon Press, Oxford, 1964)

[18] Carl M. BENDER et Stefan BOETTCHER, *"Real Spectra in Non-Hermitian Hamiltonians Having $\mathcal{PT}$ Symmetry"*, Phys. Rev. Lett., **80**, 5243 (1998)

[19] Carl M. BENDER, *"Introduction to $\mathcal{PT}$- Symmetric Quantum Theory"*, `arXiv: quant-ph/0501052`

[20] Germán SIERRA, *"A quantum mechanical model model of the Riemann zeros"*, `arXiv:http://arxiv.org/list/math-ph/0712.0705`

[21] Germán SIERRA, *"$H = xp$ with interaction and the Riemann zeros"*, Nucl. Phys. **B 327** (2007)

[22] *"Proposed proofs of the Riemann Hypothesis"* disponible à :
`www.secamlocal.ex.ac.uk/people/staff/mrwatkin/zeta/RHproofs.htm`

[23] Carl M. BENDER, *"Must a Hamiltonian be Hermitian ?"*, Am. J. Phys., **71**, 1095 (2003)

[24] Nigel GOLDENFELD, *Lectures on Phase Transitions and the Renormalization Group* (Addison - Wesley, New York, 1997)

[25] Claude COHEN-TANNOUDJI, Bernard DIU et Franck LALOË, *Mécanique Quantique* (Hermann, Paris, 1977)

[26] Roman W. JACKIW, *"Quantum meaning of classical field theory"*, Rev. Mod. Phys, **49**, 681 (1977)

[27] D. JUDGE et J.T. LEWIS, *"On the commutator $[L_z, \phi]$"*, Phys. Lett., **5**, 190 (1963)

[28] Jean-Marc LÉVY-LEBLOND, *"Who is Afraid of Nonhermitian Operators ? A Quantum Description of Angle and Phase"*, Ann. of Phys., **101**, 319 (1976)

[29] Eugen MERBACHER, *"Single Valuedness of Wave Functions"*, Am. J. Phy., **30**, 237 (1962)

[30] Richard COURANT et David HILBERT, *Methods of Mathematical Physics*, vol. I, (Interscience Pub., New York, 1953)

[31] Ernest William HOBSON, *The theory of spherical and ellipsoidal harmonic* (Cambridge University Press, 1931)

[32] Richard BARAKAT et Elizabeth BAUMANN, *"Mth Power of an $N \times N$ Matrix and its Connection with the Generalized Lucas Polynomials"*, J. Math. Phys., **10**, 1474 (1969)

[33] A. R. EDMONDS, *Angular Momentum in Quantum Mechanics*, (Princeton University Press, Princeton, 1996)

[34] Albert MESSIAH, *Mécanique Quantique* tome 1 (Dunod, Paris 1965)

[35] Franck H. STILLINGER et David R. HERRICKS, *"Bound states in the continuum"*, Phys. Rev. A, **11**, 446 (1975).

[36] John VON NEUMANN et Eugen Paul WIGNER, *"Über merkwürdige diskrete Eigenwerte"*, Physikalische Zeitschrift, **30**, 465 (1929)

[37] Milton ABRAMOWITZ et Irene A. STEGUN, *Handbook of Mathematical Functions* (Dover, New York, 1970)

[38] I.S. GRADSTEYN et I.M. RYZHIK, *Table of Integrals, Series and Products* (Academic Press, New York, 1980)

[39] Siegfried FLÜGGE, *Practical Quantum Mechanics* (Springer, Berlin, 1974)

[40] Claude ASLANGUL, *"δ well with a reflecting barrier"*, Am. J. Phys., **63**, 935 (1995)

[41] Roman W. JACKIW, *"Delta function potentials in two- and three-dimensional quantum mechanics"*, in *M. A. Bég Memorial Volume*, M. C. Gutzwiller, A. Inomata, J. R. Klauder et L. Steit éditeurs (World Scientific, Singapour, 1991)

[42] F.A.B. COUTINHO et Y. NOGAMI, *"Zero-range potential for the Dirac equation in two and three space dimensions : Elementary proof of the Svendsen's theorem"*, Phys. Rev. A, **42**, 5716 (1990)

[43] J. FERNANDO PEREZ et F.A.B. COUTINHO, *"Schrödinger equation in two dimensions for a zero-range potential and a uniform magnetic field : An exactly solvable model"*, Am. J. Phys., **59**, 52, 1991

[44] Wolfgang PAULI, *"Über das Wasserstoffspektrum vorm Standpunkt der neuen Quantenmechanik"*, Zeitschrift für Physik, **36**, 336 (1926)

[45] Vladimir FOCK, *"Zur Theorie des Wasserstoffatoms"*, Zeitschrift für Physik, **98**, 145 (1935)

[46] Harold V. McINTOSH, *"On Accidental Degeneracy in Classical and Quantum Mechanics"*, Am. J. Phys., **27**, 620, (1959)

[47] Myron BANDER et Claude ITZYKSON, *"Group Theory and the Hydrogen Atom (I)"*, Rev. Mod. Phys., **38**, 330, (1966)

[48] Lev LANDAU et Evguéni LIFSHITZ, *Mécanique Quantique* (Mir, Moscou, 1967)

[49] Edward Charles TITCHMARSH, *Theory of functions* (Oxford University Press, New York, 1988)

[50] Henri BACRY, *Leçons sur la théorie des groupes et les symétries des particules élémentaires* (Gordon et Breach, New York, 1967)

[51] Herbert GOLDSTEIN, *Mécanique classique* (Presses Universitaires de France, Paris, 1964)

[52] Valentine BARGMANN, *"Zur Theorie des Wasserstoffatoms"*, Zeitschrift für Physik, **99**, 576 (1936)

[53] Leonard I. SCHIFF, *Quantum Mechanics* (McGraw - Hill, New York, 1968)

[54] Laurent SCHWARTZ, *Méthodes mathématiques pour les sciences physiques* (Hermann, Paris, 1968)

[55] John W. DETTMAN, *Applied Complex Variables* (Macmillan, New York, 1965)

[56] Claude COHEN-TANNOUDJI, *Cours du Collège de France, 1973-1974*

[57] Lev LANDAU et Evguéni LIFSHITZ, *Théorie quantique relativiste* (Mir, Moscou, 1966)

[58] Elliott H. LIEB, *The stability of matter : from atoms to stars*, édité par Walter THIRRING (Springer, Berlin, 2001)

[59] Llewellyn THOMAS, *"The motion of a spinning electron"*, Nature, **117**, 514 (1926).

[60] Alan SOKAL et Jean BRICMONT, *Impostures intellectuelles* (Odile Jacob, Paris, 1997)

[61] `http://physics.nist.gov/cgi-bin/cuu/Value?gp|searchfor=atomnuc!`

[62] `http://physics.nist.gov/cgi-bin/cuu/Value?gnn|searchfor=atomnuc!`

[63] Satyendranath N. BOSE, *"Plancks Gesetz und Lichtquantenhypothese"*, Zeitschrift für Physik, **26**, 178 (1924).

[64] Emilio SEGRÈ, *Enrico Fermi, physicist*, University of Chicago Press (1995)

[65] Wolfgang PAULI, *"The connection between spin and statistics"*, Phys. Rev, **58**, 716 (1940)

[66] Paul O'HARA, *"The spin-statistics theorem – did Pauli get it right ?"*, disponible à `arXiv:quant-ph/0109137`

[67] Erwin SCHRÖDINGER, *Mémoires sur la Mécanique Ondulatoire* (J. Gabay, Paris, 1988)

[68] Leslie E. BALLENTINE, *Quantum Mechanics* (Prentice - Hall, Englewood Cliffs, 1990)

[69] A.W. OVERHAUSER et R. COLELLA, *"Experimental Test of Gravitationnally Induced Quantum Interference"*, Phys. Rev. Lett., **34**, 1472 (1975)

[70] R. COLELLA, A.W. OVERHAUSER et S.A. WERNER, *"Observation of Gravitationnally Induced Quantum Interference"*, Phys. Rev. Lett., **33**, 1237 (1974)

[71] H. RAUCH, A. ZEILINGER, G. BADUREK, A. WILFING, W. BAUSPIESS et U. BONSE, *"Verification of coherent spinor rotation of fermions"*, Phys. Lett., **54A**, 425 (1975)

[72] Paul Adrien Marie DIRAC, *The Development of Quantum Mechanics* (Gordon and Breach, New York, 1971)

[73] Oskar KLEIN, *"Quantentheorie und fünfdimensionale Relativitätstheorie"*, Zeitschrift für Physik, **37**, 895 (1926)

[74] Vladimir FOCK, *"Zur Schrödingerschen Wellenmechanik"*, Zeitschrift für Physik, **38**, 242 (1926)

[75] Walter GORDON, *"Der Comptoneffekt nach der Schrödingerschen Theorie"*, Zeitschrift für Physik, **40**, 117 (1926)

[76] Paul Adrien Maurice DIRAC, *"The Quantum theory of electron I"*, Proc. Roy. Soc. London, **A117**, 610 (1928) ; *"The Quantum theory of electron II"*, *ibid.* **A118**, 351 (1928)

[77] Claude ITZYKSON et Jean-Bernard ZUBER, *Quantum Field Theory* (McGraw-Hill, New York 1985)

[78] Steven WEINBERG, *The quantum theory of fields* (Cambridge University Press, Cambridge, 1995)

[79] Jun John SAKURAI, *Advanced Quantum Mechanics* (Addison-Wesley Pub. Co, Reading, 1967)

[80] Aleksandr Sergeevich DAVYDOV, *Quantum Mechanics* (Pergamon Press, 1965)

[81] Claude COHEN-TANNOUDJI, *Cours du Collège de France, 1986-1987*

[82] `http://physics.nist.gov/cgi-bin/cuu/Value?mpc2mev|searchfor=atomnuc!`

[83] Carl David ANDERSON, *"The Positive Electron"*, Phys. Rev., **43**, 491 (1933)

[84] James .D. BJORKEN et Sidney D. DRELL, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964)

[85] Charles Galton DARWIN, *"On some orbits of an electron"*, Philos. Mag., **25**, 201 (1913)

[86] Charles Galton DARWIN, *"The wave equations of the electron"*, Proc. Roy. Soc. London, **A118**, 654 (1928)

[87] Walter GORDON, *"Die Energieniveaus des Wasserstoffatoms nach der Diracschen Quantentheorie"*, Zeitschrift für Physik, **48**, 11 (1928)

[88] R.P. MARTINEZ-Y-ROMERO, *"Relativistic hydrogen atom revisisted"*, Am. J. Phys., **68**, 1050 (2000)

[89] Hans Albrecht BETHE et Edwin Ernest SALPETER, *Quantum Mechanics of one- and two-electron atoms* (Plenum-Rosetta, New York, 1977)

[90] Willis Eugene LAMB, *"Capture of Neutrons by Atoms in a Crystal"*, Phys. Rev., **55**, 190 (1939)

[91] D. T. WILKINSON et H. R. CRANE, *"Precision Measurement of the g Factor of the Free Electron"*, Phys. Rev., **130**, 852 (1963)

[92] Leslie L. FOLDY et Siegfried A. WOUTHUYSEN, *"On the Dirac Theory of Spin 1/2 Particles and Its Non-Relativistic Limit "*, Phys. Rev., **78**, 29 (1958)

[93] `http://www.britannica.com/bcom/eb/article/5/0,5716,59835+1+58374,00`

[94] B. MISRA et Ennackal Chandy George SUDARSHAN, *"The Zeno's paradox in quantum theory"*, J. Math. Phys., **18**, 756 (1977)

[95] L. A. KHALFIN, *"Contribution to the decay theory of a quasi-stationary state"*, Zh. Eksp. Teor. Fiz., **33**, 1371 (1957) [Sov. Phys. JETP, **6**, 1053 (1958)]

[96] Asher PERES, *"Zeno paradox in quantum theory"*, Am. J. Phys., **48**, 931 (1980)

[97] R. J. COOK, Phys. Scr., **T21**, 49, (1988) (cité en [98])

[98] Wayne M. ITANO, D.J. HEINZEN, J.J. BOLLINGER et D.J. WINELAND, *"Quantum Zeno effect"*, Phys. Rev. A, **41**, 2295 (1990)

[99] Bernard DIU, Claudine GUTHMANN, Danièle LEDERER et Bernard ROULET, *Physique statistique* (Hermann, Paris, 1989)

[100] Albert RENYI, *Calcul des probabilités* (Dunod, Paris, 1966)

[101] Max BORN, *"Dans quelle mesure la Mécanique classique peut-elle prédire les trajectoires ?"*, J. de Physique et le Radium, **20**, 577 (1959)

[102] Claude COHEN-TANNOUDJI, *Cours du Collège de France, 1989-1990*

[103] Anthony J. LEGGETT, Sudip CHAKRAVARTY, Alan T. DORSEY, Matthew P.A. FISCHER, Anupam GARG et Wilhelm ZWERGER, *"Dynamics of the dissipative two-state system"*, Rev. Mod. Phys., **59**, 1, (1987)

[104] Richard P. FEYNMAN et Franck Lee VERNON, Jr, *"The Theory of a General Quantum System Interacting with a Linear Dissipative System"*, Ann. Phys. (N.Y.), **24**, 118 (1963)

[105] John VON NEUMANN, *Fondements mathématiques de la Mécanique quantique* (J. Gabay, Paris, 1992), *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton 1955)

[106] Leslie E. BALLENTINE, *"Comment on* "Quantum Zeno effect" *"*, Phys. Rev. A, **43**, 5165 (1991)

[107] Almut BEIGE et Gerhardt C. HEGERFELDT, *"Projection postulate and atomic quantum Zeno effect"*, Phys. Rev. A, **53**, 53 (1996)

[108] Martin B. PLENIO et Peter L. KNIGHT, *"The quantum-jump approach to dissipative dynamics in quantum optics"*, Rev. Mod. Phys., **70**, 101 (1998)

[109] Harry PAUL, *Introduction to Quantum Theory* (Cambridge University Press, Cambridge, 2008)

[110] Th. SAUTER, W. NEUHAUSER, R. BLATT et P.E. TOSCHEK, *"Observation of Quantum Jumps"*, Phys. Rev. Lett., **57**, 1696 (1986)

[111] Hans G. DEHMELT, *"Proposed* $10^{14}\Delta\nu < \nu$ *Laser Fluorescence Spectroscopy on Tl+ Mono-Ion Oscillator II (spontaneous quantum jumps)"*, Bull. Am. Phys. Soc., **20**, 60 (1975)

[112] Paul MANNEVILLE, *Instabilités, chaos et turbulence* (Éditions de l'École Polytechnique, 2004)

[113] Richard J. COOK et H. J. KIMBLE, *"Possibility of Direct Observation of Quantum Jumps"*, Phys. Rev. Lett., **54**, 1023 (1985)

[114] Albert EINSTEIN, *"Zur Quantentheorie der Strahlung"*, Physik Zeitschrift, **18**, 121 (1917)

[115] Casey R. GRAY, *"An Analysis of the Belouzov-Zhabotinskii Reaction"*, disponible à `www.rose-hulman.edu/mathjournal/ archives/2002/vol3-n1/paper1/v3n1-1`

[116] Nicholas G. VAN KAMPEN, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 2007)

[117] Peter ZOLLER, M. MARTE et Dan F. WALLS, *"Quantum jumps in atomic systems"*, Phys. Rev., **35**, 198 (1987)

[118] Warren NAGOURNEY, Jon SANDBERG et Hans G. DEHMELT, *"Shelved Optical Electron Amplifier : Observation of Quantum Jumps"*, Phys. Rev. Lett., **56**, 2797 (1986)

[119] Jean DALIBARD, Yvan CASTIN et Klaus MØLMER, *"Wave-Function Approach to Dissipative Processes in Quantum Optics"*, Phys. Rev. Lett., **68**, 580 (1992)

[120] `http://www.cs.mcgill.ca/~crepeau/CRYPTO/Biblio-QC.html`

[121] Nicolas GISIN, Grégoire RIBORDY, Wolfgang TITTEL et Hugo ZBINDEN, *"Quantum Cryptography"*, Rev. Mod. Phys., **74**, 145 (2002)

[122] Ronald L. RIVEST, MIT Laboratory for Computer Science, Technical Memo LCS-TM82, Cambridge, Massachusetts, 4/4/77

[123] Charles H. BENNETT et Gilles BRASSARD *"Quantum Cryptography : Public Key Distribution and Coin Tossing"*, Proceedings of IEEE International Conference on Computers Systems and Signal Processing, Bangalore India, December 1984, 175-179.

[124] Artur K. EKERT, *"Quantum Cryptography Based on Bell's Theorem"*, Phys. Rev. Lett., **67**, 661 (1991)

[125] Artur K. EKERT, *"La Mécanique Quantique au secours des agents secrets"*, La Recherche, **22**, 790 (1991)

[126] Albert EINSTEIN, Boris PODOLSKI, Nathan ROSEN, *"Can Quantum Mechanical Description of Physical Reality Be Considered Complete ?"*, Phys. Rev., **47**, 777 (1935)

[127] Max JAMMER, *"Le paradoxe d'Einstein - Podolsky - Rosen"*, La Recherche, **111**, 510 (1980)

[128] Franck LALOË, *"Les surprenantes prévisions de la Mécanique quantique"*, La Recherche (novembre 1986)

[129] William Kent WOOTTERS et Wojciech Hubert ZUREK, *"A Single Quantum Cannot be Cloned"*, Nature, **299**, 802 (1982)

[130] Dennis DIEKS, *"Communication by EPR devices"*, Phys. Lett. A, **92**, 271 (1982)

[131] Michel BRUNE et Jean-Michel RAIMOND, *"L'ordinateur quantique : un défi pour les expérimentateurs"*, Images de la Physique (Éditions du CNRS, Paris, 2005)

[132] Louis DE BROGLIE, *"La thermodynamique "cachée" des particules"*, Annales de l'Institut Henri Poincaré (A) Physique théorique, **1**, p. 1 (1964)

[133] Roland OMNÈS, *Les indispensables de la Mécanique quantique* (Odile Jacob, Paris, 2007)

[134] David BOHM, *Quantum Theory* (Prentice Hall, Englewood, 1966)

[135] Hans Dieter ZEH, *"On the interpretation of measurement in quantum theory"*, Found. of Phys., **1**, 69 (1970)

[136] E. JOOS, H.D. ZEH, C. KIEFER, D. GIULINI, J. KUPSCH et I.O. STAMANESCU (éditeurs), *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin, 2007)

[137] William FELLER, *An introduction to probability theory and its applications*, Vol. I (J. Wiley, New York, 1971)

[138] Dan F. WALLS et G. J. MILBURN, *"Effect of dissipation on quantum coherence"*, Phys. Rev. A, **31**, 2403 (1985)

[139] Jean PERRIN, *Les atomes* (Flammarion, Paris, 1991)

[140] A. EINSTEIN, *"Über die von der molekular-kinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen"*, Ann. Phys. (Leipzig), **17**, 756 (1905)

[141] Lawrence S. SCHULMAN, *Techniques and Applications of Path Integration* (John Wiley, New York, 1981)

[142] Wojciech Hubert ZUREK, *"Decoherence and the Transition from Quantum to Classical* – Revisited", Los Alamos Science, **27**, 2 (2002)

[143] Hans PRIMAS, *"Zur Theorie grosser Molekeln. I. Revision der Grundlagen der Quantenchemie"*, Helv. Chim. Acta, **51**, 1037 (1968)

[144] Hans PRIMAS, *"Chemistry, quantum mechanics and reductionism"*, Lecture Notes in Chemistry, **24** (Springer, Berlin 1981)

[145] Peter PFEIFER, *"Chiral molecules - a superselection rule induced by the radiation field"*, Thesis, ETH- Zürich No. 6551 (1980)

[146] Amir Ordacgi CALDEIRA et Anthony James LEGGETT, *"Quantum Tunneling in a Dissipative System"*, Ann. Phys. (N.Y.) **149**, 374 (1983)

[147] Claude ASLANGUL, Noëlle POTTIER et Daniel Saint-JAMES, *"Spin-boson systems : equivalence between the dilute-blip and the Born approximations"*, J. Physique, **47**, 1657 (1986)

[148] Eugen Paul WIGNER, *"On the quantum correction for thermodynamic equilibrium"*, Phys. Rev., **40**, 749 (1932)

[149] Michel BRUNE, Serge HAROCHE, Jean-Michel RAIMOND, L. DAVIDOVICH et N. ZAGURY, *"Manipulation of photons in a cavity by dispersive atom-field coupling : Quantum-nondemolition measurements and generation of "Schrödinger cat" states"*, Phys. Rev. A, **45**, 5193 (1992)

[150] Erich JOOS et Hans Dieter ZEH, *"The emergence of classical properties through interaction with the environment"*, Zeitschrift für Physik B, **59**, 223 (1985)

[151] Juan Pablo PAZ, Salman HABIB et Wojciech Hubert ZUREK, *"Reduction of the wave packet : Preferred observable and decoherence time scale"*, Phys. Rev. D, **47**, 488 (1993)

[152] Michel BRUNE, E. HAGLEY, J. DREYER, Xavier MAÎTRE, Abdelhamid MAALI, C. WUNDERLICH, Jean-Michel RAIMOND et Serge HAROCHE *"Observing the Progressive Decoherence of the "Meter" in a Quantum Measurement"* , Phys. Rev. Lett., **77**, 4887 (1996)

[153] Alain ASPECT, Jean DALIBARD et Gérard ROGER, *"Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperimen : A New Violation of Bell's Inequalities"*, Phys. Rev. Lett., **49**, 91 (1982)

[154] Alain ASPECT, dans *Le Monde Quantique*, édité par Stéphane DELIGEORGE (Points Sciences, Seuil, 1984)

[155] John S. BELL, *"On the Einstein-Poldolsky-Rosen paradox"*, Physics, **1**, 195 (1964)

[156] J.F. CLAUSER, M.A. HORNE, A. SHIMONY et R.A. HOLT, *"Proposed experiment to test local hidden-variable theories"*, Phys. Rev. Lett., **23**, 880 (1969)

[157] Charles H. BENNETT, Gilles BRASSARD, C. CRÉPEAU, R. JOZSA, Asher PERES, William Kent WOOTTERS, *"Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels"*, Phys. Rev. Lett., **70**, 1895 (1993)

[158] N. David MERMIN, *"Hidden variables and the two theorems of John Bell ?"*, Rev. Mod. Phys., **65**, 803 (1993)

[159] Alain ASPECT et Philippe GRANGIER, *"De l'article d'Einstein, Podolsky et Rosen à l'information quantique : les stupéfiantes propriétés de l'intrication"*, dans *Einstein aujourd'hui*, p. 39 (EDP Sciences, CNRS Éditions, Paris, 2005)

[160] N. David MERMIN, *"Non-local character of quantum theory ?"*, `arXiv : quant-ph/9711052`

[161] N. David MERMIN, *"What Do These Correlations Know About Reality ? Nonlocality and the Absurd"*, `arXiv:quant-ph/9807055`

[162] Jean-Michel RAIMOND, Michel BRUNE et Serge HAROCHE, *Manipulating quantum entanglement with atoms and photons in a cavity"*, Rev. Mod. Phys., **73**, 565 (2001)

[163] `http://www.lkb.ens.fr/recherche/qedcav/french/rydberg/`

[164] Werner HEISENBERG, *"Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen"*, Zeitschrift für Physik, **33**, 879 (1925)

[165] Richard P. FEYNMAN, *Six Easy Pieces : Essentials of Physics Explained by Its Most Brilliant Teacher* (Penguin Books, Londres, 1998).

[166] Julius SCHWINGER, *"Spin, statistics, and the TCP theorem"*, Proc. Nat. Acad. Sci., **44**, 223 (1958) ; *ibid.* p. 617

[167] J.M. LEINAAS et J. MYRHEIM, *"On the theory of identical particles"*, Nuovo Cimento, **B37**, 1 (1977)

[168] Frank WILCZEK, *"Quantum Mechanics of Fractional-Spin Particles"*, Phys. Rev. Lett., **49**, 957 (1982)

[169] Robert Betts LAUGHLIN, *"Anomalous Quantum Hall Effect : An Incompressible Quantum Fluid with Fractionally Charged Excitations"*, Phys. Rev. Lett., **50**, 1395 (1983)

[170] L. SAMINADAYAR, D.C. GLATTLI, Y. JIN et B. ETIENNE, *"Observation of the e/3 Fractionally Charged Laughlin Quasiparticle"*, Phys. Rev. Lett., **79**, 2526 (1997)

[171] F.E. CAMINO, Wei ZHOU, V.J. GOLDMAN, *"Realization of a Laughlin quasiparticle interferometer : Observation of fractional statistics"*, Phys. Rev. B, **72**, 075342 (2005)

[172] D.E. FELDMAN et Alexei KITAEV, *"Detecting Non-Abelian Statistics with an Electronic Mach-Zehnder Interferometer"*, Phys. Rev. Lett., **97**, 166803 (2006)

[173] Daniel Edwin RUTHERFORD, *Substitutional Analysis* (Edinburgh University Press, Edinburgh, 1948)

[174] Alexander L. FETTER et John Dirk WALECKA, *Quantum Theory of Many-Particle Systems* (Dover Publications, Mineola, 2003)

[175] Erwin SCHRÖDINGER, *"Quantizierung als Eigenwertproblem (Erste Mitteilung)"*, Annalen der Physik, **79**, 361 (1926)

[176] Pierre HOHENBERG et Walther KOHN, *"Inhomogeneous Electron Gas"*, Phys. Rev, **136**, B864 (1964)

[177] Michel GAUDIN, *La fonction d'onde de Bethe* (Masson, Paris, 1983)

[178] Karl Fritiof SUNDMAN, *"Mémoire sur le problème de trois corps"*, Acta Mathematica, **36**, 105 (1912)

[179] Malte HENKEL, *"Sur la solution de Sundman du problème des trois corps"*, Philosophia Scientiae, **5**, 161-184 (2001), `arXiv:physics/0203001v1`

[180] Norman Henry MARCH, William Henry YOUNG et S. SAMPANTHAR, *The Many-Body problem in Quantum Mechanics* (Cambridge University Pres, 1967)

[181] R.D. MATTUCK, *A Guide to Feynman Diagrams in the Many-Body Problem* (McGraw-Hill, New York, 1967)

[182] Henri POINCARÉ, *Les méthodes nouvelles de la mécanique céleste* (Gauthier-Villars, Paris, 1892)

[183] Carl M. BENDER et Steven A. ORSZAG, *Advanced Mathematical Methods for Scientists and Engineers* (MacGraw-Hill, 1978)

[184] Kôsaku YOSIDA, *Équations différentielles et intégrales* (Dunod, Paris, 1971)

[185] Erwin SCHRÖDINGER, *"Quantisierung als Eigenwertproblem"*, Annalen der Physik, vol. 80, 437-490 (1926)

[186] Alexander DALGARNO et John Trevor LEWIS, *"The exact calculation of long-range forces between atoms by perturbation theory"*, Proc. Roy. Soc. London, **A 233**, 70 (1955)

[187] Paul S. EPSTEIN, *"The Stark Effect from the Point of View of Schroedinger's Quantum Theory"*, Phys. Rev., **28**, 695 (1926)

[188] Y. NOGAMI, M. VALLIÈRES et W. VAN DIJK, *"Hartree-Fock approximation for the one-dimensional helium atom"*, Am. J. Phys., **44**, 886 (1976)

[189] D. KIANG, *"Contribution of the continuum in perturbation theory"*, Am. J. Phys., **45**, 308 (1977)

[190] Mark P. SILVERMAN, *"An exactly soluble quantum model of an atom in a arbitrarily strong uniform magnetic field"*, Am. J. Phys., **49**, 546 (1981)

[191] Richard Phillips FEYNMAN, *"An Operator Calculus Having Applications in Quantum Electrodynamics"*, Phys. Rev., **84**, 108 (1951)

[192] Philip PECHUKAS et John C. LIGHT, *"On the Exponential Form of Time-Displacement Operators in Quantum Mechanics"*, J. Chem. Phys., **44**, 3897 (1966)

[193] R.M. WILCOX, *"Exponential Operators and Parameter Differentiation in Quantum Physics"*, J. Math. Phys., **8**, 962 (1967)

[194] James WEI et Edward NORMAN, *"Lie Algebraic Solution of Linear Differential Equations"*, J. Math. Phys., **6**, 1928 (1965)

[195] John D. JACKSON, *Classical Electrodynamics* (Wiley, New York, 1975)

[196] Murray GELL-MANN et Francis LOW, *"Bound States in Quantum Field Theory"*, Phys. Rev., **84**, 350 (1951)

[197] Luca Guido MOLIARI, *"Another proof of Gell-Mann and Low's Theorem"* arXiv:math-ph\0612030, J. Math. Phys., **48**, 052113 (2007)

[198] Axel WEBER, *"Relativistic bound states in Yukawa theory"*, J. Phys.: Conf. Ser. **37**, 135 (2006)

[199] Claude COHEN-TANNOUDJI, Jacques DUPONT-ROC et Gilbert GRYNBERG, *Photons et atomes : une introduction à l'électrodynamique quantique* (InterÉditions, Éditions du CNRS, Paris, 1987)

[200] Claude COHEN-TANNOUDJI, *Cours du Collège de France, 1974-1975*

[201] Claude COHEN-TANNOUDJI, *Cours du Collège de France, 1973-1974*

[202] Lev LANDAU et Evguéni LIFSHITZ, *Théorie du champ* (Mir, Moscou, 1966)

[203] Claude COHEN-TANNOUDJI, *Cours du Collège de France, 1975-1976*

[204] Claude COHEN-TANNOUDJI et Paul AVAN, *"Discrete state coupled to a continuum. Cqntinuous transition between the Weisskopf - Wigner exponential decay and the Rabi oscillation"*, Colloques Internationaux du CNRS, n° 273, p. 93 (1977), reproduit dans : Claude COHEN-TANNOUDJI, *Atoms in Electromagnetic Fields*, 2e édition, p. 255 (World Scientific, Singapour, 2004)

[205] Michel ORRIT, Claude ASLANGUL et Philémon KOTTIS, *"Quantum-mechanical-model calculations of radiative properties of a molecular crystal : I. Polaritons and abnormal decays of excitons in one- and two-dimensional systems"*, Phys. Rev., **25**, 7263 (1982)

[206] C.L. HAMMER, T.A. WEBER et V.S. ZIDELL, *"Time-dependent scattering of wave packets in one dimension"*, Am. J. Phys., **45**, 933 (1977)

[207] Paul ROMAN, *Advanced Quantum Theory* (Addison-Wesley Pub. Co., Reading, Mass. 1965)

[208] Amnon KATZ, *"Scattering Cross section for a Beam of Wave Packets"*, J. Math. Phys., **7**, 1802 (1966)

[209] Walter GORDON, *"Über den Stoßzweier Punktladungen nach der Wellenmechanik"*, Zeitschrift für Physik, **48**, 180 (1928)

[210] Alfred LANDÉ, *"Termstruktur und Zeemaneffekt der Multipletts"*, Zeitschrift für Physik, **15**, 189 (1923)

[211] Egil A. HYLLERAAS, *"Neue berechnung der energie des helium im grundzustande sowie des tiefsten terms von Ortho-Helium "*, Zeitschrift für Physik, **54**, 347 (1929)

[212] Harrison SHULL et Per-Olov LÖWDIN, *"Role of the Continuum in Superposition of Configurations"*, J. Chem. Phys., **23**, 1362 (1955)

[213] Robert T. BROWN, *"Electronic Interaction Integrals for Atoms Calculated with Laguerre Polynomial Radial Wavefunctions"*, J. Chem. Phys., **46**, 1551 (1967)

[214] M.K. ODLING - SMEE, E. SOKELL, P. HAMMOND et M.A. MacDONALD, *"Radiative Decay of Doubly Excited States in Helium below the* He^+ $(N = 2)$ *Ionization Threshold"*, Phys. Rev. Lett., **84**, 2598 (2000)

[215] John Clarke SLATER, *Quantum Theory of Molecules and Solids*, vol. I (McGraw-Hill, New York, 1963)

[216] Jürg FRÖLICH, *"Spin or, Actually : Spin and Quantum Statistics"*, Séminaire Poincaré XI (2007), p. 1

[217] Robert Nyden HILL, *"Proof that the* H^- *Ion Has Only One Bound State"*, Phys. Rev. Lett., **38**, 643 (1977)

[218] Hans Albrecht BETHE, *"Berechnung der Elektronenaffinität des Wasserstoffs"*, Zeitschrift für Physik, **57**, 815 (1929)

[219] Oeyvind BURRAU, *"Berechnung des Energiewertes des Wasserstoff Molekel-Ions* (H_2^+) *im Normalzustand"*, Kgl. Danske Vid. Selsk. Matt.-Fys. Medd., **7**, 14 (1927)

[220] I. N. LEVINE, *Quantum Chemistry*, vol. I (Allyn & Bacon, Boston, 1970)

[221] Walter Heinrich HEITLER et Fritz Wolfgang LONDON, *"Wechselwirkung neutraler Atome und Homöopolare Bindung nach der Quantenmechanik"*, Zeitschrift für Physik, **44**, 455 (1927)

[222] Giorgina CORONGIU, *"The Hartree - Fock - Heitler - London Method, III : Correlated Diatomic Hydrides"*, J. Phys. Chem A, **111**, 5333 (2007)

[223] S.C. WANG, *"The Problem of the Normal Hydrogen Molecule in the New Quantum Mechanics"*, Phys. Rev., **31**, 579 (1928)

[224] N. ROSEN, *"The Normal State of the Hydrogen Molecule"*, Phys. Rev., **38**, 2099 (1931)

[225] Erich HÜCKEL, *"Grundzüge der Theorie Ungesättigter und Aromatischer Verbindungen"*, Zeitschrift für Physik, **70**, 204, (1931)

[226] C.K. CHIANG, M.A. DRUY, S.C. GAN, A.J. HEEGER, E.J. LOUIS, A.G. Mac DIARMID, Y.W. PARK et H. SHIRAKAWA, *"Synthesis of highly conducting films of derivatives of a polyacetylene, $(CH)_x$"*, J. Am. Chem. Soc. **100**, 1013 (1978)

[227] Mikio TSUJI, Sigeru HUZINAGA et Tasuke HASINO, *"Bond Alternation in Long Polyenes"*, Rev. Mod. Phys., **32**, 425 (1960)

[228] W.P. SU, J.R. SCHRIEFFER et A.J. HEEGER, *"Soliton excitations in polyacetylene"*, Phys. Rev. B, **22**, 2099 (1980)

[229] Claude ASLANGUL et Daniel SAINT-JAMES, *"Continuum Hartree - Fock scheme for 1-D polymers"*, J. de Phys., **45**, 1401 (1984)

[230] Michael TINKHAM, *Group Theory in Quantum Mechanics* (New York , McGraw-Hill, 1964)

[231] Federick REIF, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, 1987)

[232] Eugen Paul WIGNER, *"On the Interaction of Electrons in Metals"*, Phys. Rev., **46**, 1002 (1934)

[233] C.C. GRIMES et G. ADAMS, *"Evidence for a Liquid-to-Crystal Phase Transition in a Classical Two-Dimensional Sheet of Electrons"*, Phys. Rev. Lett., **42**, 795 (1979)

[234] D. SHECHTMAN, I. BLECH, D. GRATIAS et J. W. CAHN, *"Metallic Phase with Long-Range Orientational Order and No Translational Symmetry"* Phys. Rev. Lett., **53**, 1951 (1984)

[235] W. MARSHALL et S.W. LOVESEY, *Theory of Thermal Neutron Scattering* (Holt, Rinehart et Winston, Oxford, 1971)

[236] N.W. ASHCROFT et N. David MERMIN, *Solid State Physics* (Clarendon Press, Oxford, New York 1976)

[237] Markus GREINER, Olaf MANDEL, Tilman ESSLINGER, Theodor W. HÄNSCH and Immanuel BLOCH, *"Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms"*, Nature **415**, 39 (2002)

[238] J. LÉONARD, M. WALHOUT, A.P. MOSK, T. MÜLLER, Michèle LEDUC et Claude COHEN-TANNOUDJI, *"Giant Helium Dimers Produced by Photoassociation of Ultracold Metastable Atoms"*, Phys. Rev. Lett., **91**, 073203 (2003)

[239] William C. STWALLEY, Yea-Hwang UANG et Goran PICHLER *"Pure Long-Range Molecules"*, Phys. Rev. Lett., **41**, 1164 (1978)

[240] J. HUBBARD, *"Electron Correlations in Narrow Energy Bands"*, Proc. Roy. Soc. London, **A276**, 238 (1963)

[241] Elliott H. LIEB et F.Y. WU, *"Absence of Mott transition in an exact solution of the short-range one-band model in one dimension"*, Phys. Rev. Lett., **20**, 1445 (1968); erratum, ibid. **21**, 192 (1968).

[242] Elliott H. LIEB et F.Y. WU, *"The one-dimensional Hubbard model : A reminiscence"*, `arXiv:cond-mat/0207529v2`

[243] Philippe NOZIÈRES et Joaquin Mazdak LUTTINGER, *"Derivation of the Landau Theory of Fermi Liquids. I. Formal Preliminaries"*, Phys. Rev., **127**, 1423 (1962)

[244] Joaquin Mazdak LUTTINGER et Philippe NOZIÈRES, *"Derivation of the Landau Theory of Fermi Liquids. II. Equilibrium Properties and Transport Equation"*, Phys. Rev., **127**, 1431 (1962)

[245] David PINES et Philippe NOZIÈRES, *The Theory of Quantum Liquids, Volume I: Normal Fermi Liquids* (W.A. Benjamin, New York, 1966), Westview Press, Boulder, 1999)

[246] `http://monet.unibas.ch/ elmer/flab/FKModel/`

[247] Hans Albrecht BETHE, *"On the theory of metals. 1. Eigenvalues and eigenfunctions for the linear atomic chain"*, Zeitschrift für Physik, **71**, 205 (1931)

[248] Philip Warren ANDERSON, *"Localized Magnetic States in Metals"*, Phys. Rev., **124**, 41 (1961)

[249] Ekkehard PEIK, Maxime BEN DAHAN, Isabelle BOUCHOULE, Yvan CASTIN et Christophe SALOMON, *"Bloch oscillations of atoms, adiabatic rapid passage, and monokinetic atomic beams"*, Phys. Rev. A, **55**, 2989 (1997)

[250] S. K. YIP, *"Dimer State of Spin - 1 Bosons in an Optical Lattice"*, Phys. Rev. Lett., **90**, 250402 (2003)

[251] Julien VIDAL, Rémy MOSSERI et Benoît DOUÇOT, *"Aharonov-Bohm Cages in Two-Dimensional Structures"*, Phys. Rev. Lett., **81**, 5888 (1998)

[252] Julien VIDAL, Benoît DOUÇOT, Rémy MOSSERI et Patrick BUTAUD, *"Interaction Induced Delocalization for Two Particles in a Periodic Potential"*, Phys. Rev. Lett., **81**, 5888 (1998)

[253] M.N. BAIBICH, J.M. BROTO, A. FERT, F. NGUYEN VAN DAU, F. PETROFF, P. EITENNE, G. CREUZET, A. FRIEDRICH et J. CHAZELAS, *"Giant Magnetoresistance of (001)Fe/(001)Cr Magnetic Superlattices"* Phys. Rev. Lett., **61**, 2472 (1988)

[254] P. GRÜNBERG, R. SCHREIBER, Y. PANG, M.B. BRODSKY et H. SOWERS (1986), *"Layered Magnetic Structures : Evidence for Antiferromagnetic Coupling of Fe Layers across Cr Interlayers"*, Phys. Rev. Lett. **57**, 2442 (1988)

[255] Eckehard SCHÖLL (éditeur), *Theory of Transport Properties of Semiconductor Nanostructures* (Chapman and Hall, Londres, 1998)

[256] Heike KAMERLINGH ONNES, *"The Superconductivity of Mercury"*, Comm. Phys. Lab. Univ. Leiden, Nos. 122 and 124 (1911)

[257] Johannes Georg BEDNORZ et Karl Alexander MÜLLER, *"Possible high Tc superconductivity in the Ba-La-Cu-O system"*, Zeitschrift für Physik B, **64**, 189 (1986)

[258] Patrick A. LEE, Naoto NAGAOSA et Xiao-Gang WEN, *"Doping a Mott insulator : Physics of high-temperature superconductivity"*, Rev. Mod. Phys., **78**, 17 (2006)

[259] Roland COMBESCOT et Xavier LEYRONAS, *"A simple theory for high* Δ/T_c *ratio in d-wave superconductors"*, Eur. Phys. J. B, **23**, 159 (2001)

[260] Elbio DAGOTTO, *"Correlated electrons in high-temperature superconductors"*, Rev. Mod. Phys., **66**, 763 (1994)

[261] Walther MEISSNER et Robert OCHSENFELD, *"Ein neuer Effekt bei eintritt der Supraleitfähigkeit"*, Naturwissenschaften, **21**, 787 (1933)

[262] Emanuel MAXWELL, *"Isotope Effect in the Superconductivity of Mercury "*, Phys. Rev., **78**, 477 (1950)

[263] William S. CORAK, B.B. GOODMAN, C.B. SATTERTHWAITE et Aaron WEXLER, *"Exponential Temperature Dependence of the Electronic Specific Heat of Superconducting Vanadium"*, Phys. Rev., **96**, 1442 (1954)

[264] Leon N. COOPER, *"Bound Electron Pairs in a Degenerate Fermi Gas"*, Phys. Rev., **104**, 1189 (1956)

[265] John BARDEEN, Leon N. COOPER et John Robert SCHRIEFFER, *"Microscopic Theory of Superconductivity"*, Phys. Rev., **106**, 162 (1957)

[266] Denis JÉROME, A. MAZAUD, M. RIBAULT et Klaus BECHGAARD, *"Superconductivity in a synthetic organic conductor :* $(TMTSF)_2PF_6$ *"*, J. Phys. Lett., **41**, L95 (1980).

[267] Michael TINKHAM, *Introduction to Superconductivity* (McGraw-Hill, New York, 1996)

[268] Herbert FRÖLICH, *"Interaction of electrons with lattice vibrations"*, Proc. Roy. Soc. London, **A215**, 291 (1952)

[269] Nikolai Nikolaevich BOGOLYUBOV, *"On a new method in the theory of superconductivity"*, Nuovo Cimento, **7**, 794 (1958)

[270] J. G. VALATIN, *"Comments on the Theory of Superconductivity"*, Nuovo Cimento, **7**, 843 (1958)

[271] Marc MÉZARD et Giorgio PARISI, *"Replica field theory for random manifolds"*, J. de Phys., **I 1**, 809 (1991)

[272] W.A. LITTLE, *"Possibility of Synthesizing an Organic Superconductor"*, Phys. Rev. A, **134**, 1416 (1964)

Index

D

E

F

G

H

I

J - K

L

M

Q

R

S

T

U

V

W

X - Y - Z

Claude **Aslangul**

Mécanique quantique 2

Ce deuxième tome s'adresse aux étudiants de Master. Dans le même état d'esprit que le premier tome, on y développe le formalisme dans des situations plus complexes, s'appuyant sur des considérations physiques et expliquant les concepts dans un langage aussi intuitif et accessible que possible. L'arsenal de mathématiques appliquées nécessaire à la maîtrise du sujet est développé au fur et à mesure, belle occasion d'introduire quelques outils indispensables au physicien, quelle que soit sa spécialité.

La première partie s'appuie notamment sur la notion de symétrie. La théorie du moment cinétique et le champ central sont exposés. L'introduction du spin est faite sur des bases physiques, conduisant à l'équation de Dirac et à sa discussion. Les postulats quantiques sont ensuite revisités à la lumière d'expériences récentes, permettant de revenir sur les étrangetés quantiques (intrication), la décohérence et des applications surprenantes (cryptographie). Après l'exposé des principes des méthodes perturbatives et variationnelles, les bases de la quantification du rayonnement sont expliquées. Cette partie se termine par une introduction à la théorie des collisions.

La deuxième partie propose quelques applications, délibérément restreintes à la physique de basse énergie, où on s'efforce de montrer l'universalité des concepts quantiques dans des champs aussi variés que la physique atomique, la chimie et la physique des solides, permettant de mettre en lumière l'immense pouvoir explicatif et les innombrables succès de la théorie quantique.

Cet ouvrage est issu d'une expérience d'enseignement pendant plusieurs années en Licence et Maîtrise de physique de l'Université Pierre et Marie Curie (Paris 6) et à l'École Normale Supérieure (Ulm).



Les «plus»

- Très nombreux exercices corrigés
- Style clair et vivant, facilitant la lecture
- Conforme aux enseignements de 4e année de physique (M1)

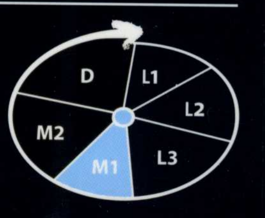
Claude Aslangul *est professeur à l'Université Pierre et Marie Curie (Paris 6) et enseigne également à l'École Normale Supérieure (Ulm). Il est membre du Laboratoire de Physique Théorique de la Matière Condensée (Jussieu).*

ISBN 978-2-8041-5881-1



MECQUADEV

Dans le cadre du nouveau Système Européen de Transfert de Crédits (E.C.T.S.), ce manuel couvre les niveaux :
En **France :** Master 1.
En **Belgique** : Master 1.
En **Suisse** : Master 1.
Au **Canada :** Master 1.



Conception graphique : Primo&Primo